This Page Is Inserted by IFW Operations and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Image Problems Mailbox.

THIS PAGE BLANK (USPTO)

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau





INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6: C12N 15/29, C07K 14/415, C12N 5/04,

5/14, A01H 5/00, C12N 15/10, 15/82

A2

(11) International Publication Number:

(43) International Publication Date:

22 April 1999 (22.04.99)

WO 99/19492

(21) International Application Number:

PCT/EP98/06977

(22) International Filing Date:

9 October 1998 (09.10.98)

(30) Priority Data:

PO 9745

10 October 1997 (10.10.97)

(71) Applicant (for all designated States except US): RHONE-POULENC AGRO [FR/FR]; 14/20, rue Pierre Baizet, F-69009 Lyon (FR).

(72) Inventors; and

- (75) Inventors/Applicants (for US only): DOUTRIAUX, Marie-Pascale [FR/FR]; 64, route de Villebon, F-91160 Saulx les Chartreux (FR). BETZNER, Andreas, Stefan [AU/AU]; 40 Dallachy Place, Page, ACT 2614 (AU). FREYSSINET, Georges [FR/FR]; 21, rue de Nervieux. F-69450 Saint Cyr au Mont d'Or (FR). PEREZ, Pascal [FR/FR]; 17, chemin de la Pradelle, Varennes, F-63450 Chanonat (FR).
- (74) Agent: GENIN, Patrick; Rhône-Poulenc Agro, DPI, 14/20, rue Pierre Baizet, F-69009 Lyon (FR).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

- (54) Title: METHODS FOR OBTAINING PLANT VARIETIES

An isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide functionally involved in the DNA mismatch repair system of a plant.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GB	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	ΙE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of Americ
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP.	Democratic People's	NZ	New Zealand	211	Zimbabwe
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	ΚZ	Kazakstan	RO	Romania		
CZ	Czech Republic	ic	Saint Lucia	RU	Russian Federation		
DE	Germany	u	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE .	Estonia	LR	Liberia	SG	Singapore		

10

Methods for Obtaining Plant Varieties

TECHNICAL FIELD

The present invention relates to nucleotide sequences which encode polypeptides involved in the DNA mismatch repair systems of plants, and to the polypeptides encoded by those nucleotide sequences. The invention also relates to nucleotide sequences and polypeptide sequences for use in altering the DNA mismatch repair system in plants. The invention also relates to a process for altering the DNA mismatch repair system of a plant cell, to a process for increasing genetic variations in plants and to processes for obtaining plants having a desired characteristic.

BACKGROUND OF THE INVENTION

Plant breeding essentially relies on and makes use of genetic variation which occurs naturally within and between members of a family, a genus, a species or a subspecies. Another source of genetic variation is the introduction of genes from other organisms which may or may not be related to the host plant.

Allelic loci or non-allelic genes which constitute or contribute to desired quantitative (e.g. growth performance, yield, etc.) or qualitative (e.g. deposition, content and composition of seed storage products; pathogen resistance genes: etc.) traits that are absent, incomplete or inefficient in a species or subspecies of interest are typically introduced by the plant breeder from other species or subspecies, or *de novo*. This introduction is often done by crossing, provided that the species to be crossed are sexually compatible. Other means of introducing genomes, individual chromosomes or genes into plant cells or plants are well known in the art. They include cell fusion, chemically aided transfection (Schocher et al., 1986, Biotechnology 4: 1093) and ballistic (McCabe et al., 1988, Biotechnology 6: 923), microinjection (Neuhaus et al., 1987, TAG 75: 30), electroporation of protoplasts (Chupeau et al., 1989, Biotechnology 7: 53) or microbial transformation methods such as Agrobacterium mediated transformation (Horsch et al., 1985, Science 227: 1229; Hiei et al., 1996, Biotechnology 14: 745).

However, when a foreign genome, chromosome or gene is introduced into a plant, it will often segregate in subsequent generations from the genome of the recipient plant or plant cell during mitotic and meiotic cell divisions and, in consequence, become lost from the host plant or plant cell into which it had been introduced. Occasionally, however, the introduced genome, chromosome or gene physically combines entirely or in part with the genome, chromosome or gene of the host plant or plant cell in a process which is called recombination.

Recombination involves the exchange of covalent linkages between DNA molecules in regions of identical or similar sequence. It is referred to here as homologous recombination if donor and recipient DNA are identical or nearly identical (at least 99%)

hase sequence identity), and as homeologous recombination if donor and recipient DNA are not identical but are similar (less than 99% base sequence identity).

The ability of two genomes, chromosomes or genes to recombine is known to depend largely on the evolutionary relation between them and thus on the degree of sequence similarity between the two DNA molecules. Whereas homologous recombination is frequently observed during mitosis and meiosis, homeologous recombination is rarely or never seen.

From a breeder's perspective, the limits within which homologous recombination occurs, therefore, define a genetic barrier between species, varieties or lines, in contrast to homeologous recombination which can break this barrier. Homeologous recombination is thus of great importance for plant breeding. Accordingly there is a need for a process for enhancing the frequency of homeologous recombination in plants. In particular, there is a need for a process of increasing homeologous recombination to significantly shorten the length of breeding programs by reducing the number of crosses required to obtain an otherwise rare recombination event.

At least in Escherichia coli, homologous and homeologous recombination are known to share a common pathway that requires among others the proteins RecA, RecB, RecC. RecD and makes use of the SOS induced RuvA and RuvB, respectively. It has been suggested that mating induced recombination follows the Double-Strand Break Repair 20 model (Szostak et al., 1983, Cell 33, 25-35), which is widely used to describe genetic recombination in eukaryotes. Following the alignment of homologous or homeologous DNA double helices the RecA protein mediates an exchange of a single DNA strand from the donor helix to the aligned recipient DNA helix. The incoming strand screens the recipient helix for sequence complementarity, seeking to form a heteroduplex by hydrogen 25 bonding the complementary strand. The displaced homologous or homeologous strand of the recipient helix is guided into the donor helix where it base pairs with its counterpart strand to form a second heteroduplex. The resulting branch point then migrates along the aligned chromosomes thereby elongating and thus stabilising the initial heteroduplexes. Single stranded gaps (if present) are closed by DNA synthesis. The strand cross overs 30 (Holliday junction) are eventually resolved enzymatically to yield the recombination products.

Although in wild type E. coli homologous and homeologous recombination are thus mechanistically similar if not identical, homologous recombination in conjugational crosses E. coli x E. coli occurs five orders of magnitude more frequently than homeologous recombination in conjugational crosses E. coli x S. typhimurium (Matic et al. 1995; Cell 80, 507-515). The imbalance in favour of homologous recombination was shown to be caused largely by the bacterial MisMatch Repair (MMR) system since its

inactivation increased the frequency of homeologous recombination in E. coli up to 1000 fold (Rayssiguier et al. 1989, Nature 342, 396-401).

In E. coli. the MMR system (reviewed by Modrich 1991, Annual Rev Genetics 25, 229-253) is composed of only three proteins known as MutS, MutL and MutH. MutS recognizes and binds to base pair mismatches. MutL then forms a stable complex with mismatch bound MutS. This protein complex now activates the MutH intrinsic single stranded endonuclease which nicks the strand containing the misplaced base and thereby prepares the template for DNA repair enzymes.

During recombination, MMR components inhibit homeologous recombination. In vitro experiments demonstrated that MutS in complex with MutL binds to mismatches at the recombination branch point and physically blocks RecA mediated strand exchange and heteroduplex formation (Worth et al., 1994; PNAS 91, 3238-3241). Interestingly, the SOS dependent RuvAB mediated branch migration is insensitive to MutS/MutL, explaining the observed slight increase in SOS dependent homeologous recombination.

15 Homeologous mating even induces the SOS response, thereby taking advantage of RuvAB induction (Matic et al. 1995, Cell 80, 507-515).

The MMR system thus appears to be a genetic guardian over genome stability in E. coli. In this role it essentially determines the extent to which genetic isolation, that is, speciation, occurs. The diminished sensitivity of the SOS system to MMR, however, allows (within limits) for rapid genomic changes times of stress, providing the means for fast adaptation to altered environmental conditions and thus contributing to intraspecies genetic variation and species evolution.

The important role of MMR in preserving genomic integrity has been established also in certain eukaryotes. In its efficiency, the human MMR, for example, may even counteract potential gene therapy tools such as triple-helix forming oligonucleotides including RNA-DNA hybrid molecules (Havre et al., 1993, J. Virology 67: 7234-7331; Wang et al., 1995, Mol. Cell. Biol. 15: 1759-1768; Kotani et al., 1996, Mol. Gen. Genetics 250: 626-634; Cole-Strauss et al., 1996, Science 273: 1387-1389). Such oligonucleotides are designed to introduce single base changes into selected DNA target sequences in order to inactivate for example cancer genes or to restore their normal function. The resulting base mismatches however are recognised by the mismatch repair system which then directs removal of the mismatched base, thereby reducing the efficiency of oligonucleotide induced site-specific mutagenesis.

To date, two families of related genes, homologous to the bacterial MutS and MutL genes have been identified or isolated in yeast and mammals (recent reviews by Arnheim and Shibata, 1997, Curr. Opinion Genet. Dev. 7, 364-370; Modrich and Lahue, 1996, Annual Rev. Biochem. 65, 101-133; Umar and Kunkel, 1996, Eur. J. Biochem. 238, 297-307). Biochemical and genetic analysis indicated that eukaryotic MutS homologs (MSH)

and MutL homologs (MLH, PMS), respectively, fulfil similar protein functions as their bacterial counterparts. Their relative abundance, however, could reflect different mismatch specificity and/or specialisation for different tissues or organelles or developmental processes such as mitotic versus meiotic recombination.

To date, six different genes homologous to *MutS* have been isolated in yeast (yMSH), and their homologs have been found in mouse (mMSH) and human (hMSH), respectively. Encoded proteins yMSH2, yMSH3 and yMSH6 appear to be the main MutS homologs involved in MMR during mitosis and meiosis in yeast, where the complementary proteins MSH3 and MSH6 alternatively associate with MSH2 to recognise different mismatch substrates (Masischky et al., 1996, Genes Dev. 10, 407-420). Similar protein interactions have been demonstrated for the human homologs hMSH2, hMSH3 and hMSH6 (Acharya et al., 1996, PNAS 93, 13629-13634).

MutL homologs (MLH and PMS), recently reviewed by Modrich and Lahue (1996, Annual Rev. Biochem. 65, 101-133) have so far been found in yeast (yMLH1 and yPMS1), mouse (mPMS2) and human (hMLH1, hPMS1 and hPMS2). The hPMS2 is a member of a family of at least 7 genes (Horii et al., 1994, Biochem. Biophys. Res. Commun. 204, 1257-1264) and its gene product is most closely related to yPMS1. Prolla et al. (1994, Science 265, 1091-1093) presented evidence for yPMS1 and yMLH1 to physically associate with each other and, together, to interact with the MutS homolog yMSH2 to form a ternary complex involved in mismatch substrate binding.

However, while medical interest in mismatch repair has prompted extensive research on MMR in bacteria, yeast and mammals, MMR genes have not been isolated from higher plants prior to the present invention and no attempts to adjust the plant MMR to plant breeding needs have been reported.

SUMMARY OF THE INVENTION

According to a first embodiment of the invention, there is provided an isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide functionally involved in the DNA mismatch repair system of a plant. In one form of this embodiment, the invention provides an isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human. More particularly, the invention provides polynucleotide sequences encoding polypeptides which are homologous to the mismatch repair polypeptides MSH3 and MSH6 of Saccharomyces cerevisiae. Still more particularly, the invention provides the coding sequences of the genes AtMSH3 and AtMSH6 of Arabidopsis thaliana. as defined hereinbelow, and polynucleotide sequences encoding polypeptides which are homologous to polypeptides encoded by AtMSH3 and AtMSH6.

25

According to a second embodiment of the invention, there is provided an isolated and purified polypeptide functionally involved in the DNA mismatch repair system of a plant, for example a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human such as a polypeptide encoded by the genes AtMSH3 or AtMSH6 of Arabidopsis thaliana, as defined hereinbelow.

5

According to a third embodiment of the invention, there is provided an isolated and purified DNA molecule comprising a polynucleotide sequence selected from the group consisting of (i) a sequence encoding a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence; and (ii) a sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant.

According to a fourth embodiment of the invention there is provided a chimeric gene comprising a DNA sequence selected from the group consisting of (i) a sequence encoding a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence, and (ii) a sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant: together with at least one regulation element capable of functioning in a plant cell. Examples of such regulation elements include constitutive, inducible, tissue type specific and cell type specific promoters such as 35S. NOS, PR1a, AoPR1 and DMC1. Typically, a chimeric gene of the fourth embodiment will also include at least one terminator sequence, more typically exactly one terminator sequence.

In the third and fourth embodiments, said interference, by said polynucleotide sequence, with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair peptide of a yeast or a human typically occurs by hybridisation or by co-suppression.

According to a fifth embodiment of the invention there is provided a plasmid or vector comprising a chimeric gene of the fourth embodiment. A vector of the fifth 30 embodiment may be, for example, a viral vector or a bacterial vector.

According to a sixth embodiment of the invention, there is provided a plant cell stably transformed, transfected or electroporated with a plasmid or vector of the fifth embodiment.

According to seventh embodiment of the invention, there is provided a plant somprising a cell of the sixth embodiment.

According to an eighth embodiment of the invention, there is provided a process for at least partially inactivating a DNA mismatch repair system of a plant cell, comprising

transforming or transfecting said plant cell with a DNA sequence of the third embodiment or a chimeric gene of the fourth embodiment or a plasmid or vector of the fifth embodiment, and causing said DNA sequence to express said polynucleotide or said polypeptide.

According to a ninth embodiment of the invention, there is provided a process for increasing genetic variation in a plant comprising obtaining a hybrid plant from a first plant and a second plant, or cells thereof, said first and second plants being genetically different; altering the mismatch repair system in said hybrid plant; permitting said hybrid plant to self-fertilise and produce offspring plants; and screening said offspring plants for plants in which homeologous recombination has occurred. For example, homeologous recombination may be evidenced by new genetic linkage of a desired characteristic trait or of a gene which contributes to a desired characteristic trait.

According to a tenth embodiment of the invention there is provided a process for obtaining a plant having a desired characteristic, comprising altering the mismatch repair system in a plant, cell or plurality of cells of a plant which does not have said desired characteristic, permitting mutations to persist in said cells to produce mutated plant cells, deriving plants from said mutated plant cells, and screening said plants for a plant having said desired characteristic.

In a preferred form of the ninth and tenth embodiments of the invention, the step of altering the mismatch repair system comprises introducing into said hybrid plant, plant, cell or cells a chimeric gene of the fourth embodiment and permitting the chimeric gene to express a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence in a mismatch repair gene of the hybrid plant, plant, cell or cells, or a polypeptide capable of disrupting the DNA mismatch repair system of the hybrid plant or cells.

In other embodiments, the invention provides (a) an oligonucleotide capable of hybridising at 45°C under standard PCR conditions to a DNA molecule of the first embodiment; (b) an oligonucleotide capable of hybridising at 45°C under standard PCR conditions to the DNA of SEQ ID NO: 18 and (c) an oligonucleotide capable of hybridising at 45°C under standard PCR conditions to the DNA of SEQ ID NO:30; with the proviso that the oligonucleotide of (a), (b) and (c) is other than SEQ ID NO:1 or SEQ ID NO:2.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 provides a diagrammatic representation of the primer sequences used to isolate AtMSH3.

Figure 2 is a plasmid map of clone 52, showing restriction enzyme cleavage sites in the 5' half of the full-length cDNA for AtMSH3.

Figure 3 is a plasmid map of clone 13, showing restriction enzyme cleavage sites in the 3' half of the full-length cDNA for AtMSH3.

Figure 4 is a sequence listing of the coding sequence of AtMSH3, together with a deduced sequence of the encoded polypeptide.

Figure 5 is a protein alignment of yeast (Saccharomyces cerevisiae) and Arabidopsis thaliana MSH3 protein.

Figure 6 provides a diagrammatic representation of the primer sequences used to isolate AtMSH6.

Figure 7 is a plasmid map of clone 43, showing restriction enzyme cleavage sites in the 5' half of the full-length cDNA for AtMSH6.

Figure 8 is a plasmid map of clone 62, showing restriction enzyme cleavage sites in the 3' half of the full-length cDNA for AtMSH6.

Figure 9 is a sequence listing of the coding sequence of AtMSH6, together with a deduced sequence of the encoded polypeptide.

Figure 10 is a protein alignment of yeast (Saccharomyces cerevisiae) and Arabidopsis thaliana MSH6 protein.

Figure 11 is a genomic sequence listing of AtMSH6.

Figure 12 is a plasmid map of plasmid pPF13.

Figure 13 is a plasmid map of plasmid pPF14.

Figure 14 is a plasmid map of plasmid pCW186.

Figure 15 is a plasmid map of plasmid pCW187.

Figure 16 is a plasmid map of plasmid pPF66.

Figure 17 is a plasmid map of plasmid pPF57.

Figure 18 is a diagrammatic representation of an antisense gene construction for use in homeologous meiotic recombination.

Figure 19 is a plasmid map of plasmid p3243.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is based on the inventors' discovery that there exist in higher plants genes which are homologous to MMR genes in E. coli, and to MMR genes in yeasts and humans.

Thus, the inventors have identified genes, herein designated AtMSH3 and AtMSH6, of the plant Arabidopsis thaliana which encode the proteins AtMSH3 and AtMSH6. These plant proteins are homologous to yMSH3 and yMSH6, respectively. The present inventors have isolated cDNAs encoding the proteins AtMSH3 and AtMSH6 and have isolated the complete gene encoding AtMSH6. Given the teaching herein, other genes (for example AtMSH2, and genes of other plants) may be obtained which are involved in DNA mismatch repair in plants, including other genes which encode polypeptides homologous to MMR proteins of yeasts or humans, such as genes which encode

20

WO 99/19492 PCT/EP98/06977

polypeptides homologous to yeast MSH2. MLH1 or PMS2, or to human MLH1. PMS1 or PMS2. For example, given the teaching herein, genes of members of the *Brassicaceae* family or of other unrelated families, for example the *Poaceae*, the *Solanaceae*, the *Asteraceae*, the *Malvaceae*, the *Fabaceae*, the *Linaceae*, the *Canabinaceae*, the *Dauaceae* and the *Cucurbitaceae* family, and which encode polypeptides homologous to MMR proteins of yeasts or humans may be obtained.

Examples of plants whose genes encoding polypeptides homologous to MMR proteins of yeasts or humans may be obtained given the teaching herein include maize, wheat, oats, barley, rice, tomato, potato, tobacco, capsicum, sunflower, lettuce, artichoke, safflower, cotton, okra, beans of many kinds including soybean, peas, melon, squash, cucumber, oilseed rape, broccoli, cauliflower, cabbage, flax, hemp, hops and carrot.

Within the meaning of the present invention, a first polypeptide is defined as homologous to a second polypeptide if the amino acid sequence of the first polypeptide exhibits a similarity of at least 50% on the polypeptide level to the amino acid sequence of the second polypeptide.

A procedure which may be followed to obtain genes AtMSH3 and AtMSH6 is described in Example 1. Essentially the same technique may be applied to obtain other mismatch repair genes of Arabidopsis thaliana, and essentially the same technique as 20 exemplified herein may be applied to cDNA obtained by reverse transcription of RNA from other plants. Alternatively, given the sequence information disclosed herein, other degenerate oligonucleotide primers, especially oligonucleotides of the invention which are capable of hybridising at 45°C under standard PCR conditions (such as the conditions described in Example 1 using primers UPMU and DOMU) to AtMSH3 and/or AtMSH6 25 may be designed and obtained for use in isolating sequences of plant mismatch repair genes which are homologous to AtMSH3 or AtMSH6, from other plants. oligonucleotides of the invention which are capable of hybridising at 45°C under standard PCR conditions to plant mismatch repair genes of plants other than Arabidopsis thaliana also fall within the scope of the present invention and may be utilised to obtain mismatch Typically, such oligonucleotides are capable of 30 repair genes of still other plants. hybridising at 45°C under standard PCR conditions to a DNA molecule which encodes a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or a human. The temperature at which oligonucleotides of the invention hybridise to AtMSH3 and/or AtMSH6, or to plant mismatch repair genes of plants other than Arabidopsis thaliana, or 35 to DNA molecules which encode polypeptides which are homologous to a mismatch repair polypeptide of a yeast or a human may be higher than 45°C, for example at least 50°C, or at least 55°C, or at least 60°C or as high as 65°C.

The successful gene isolation disclosed herein demonstrates for the first time the existence of MMR in higher plants and indicates the presence of other plant MMR genes. For example, genes encoding the plant homologs of MSH1, MSH2, MSH4, MSH5, PMS1, PMS2 and MLH1 may be identified given the teaching herein. Such genes, as well as those specifically described herein, separately or in combination, are useful in manipulating the plant MMR for plant breeding purposes. Thus, for example, the plant MMR may be altered by including in a plant cell a polynucleotide sequence as defined herein above with reference to the third embodiment of the invention, and causing the polynucleotide sequence to express either a polynucleotide which disables a plant MMR genes.

The DNA molecule of the third embodiment of the invention includes a polynucleotide sequence (herein referred to as a MMR altering gene) which may for example encode sense, antisense or ribozyme molecules characterised by sufficient base sequence similarity or complementarity to the gene to be altered to permit the antisense or 15 ribozyme molecule to hybridise with the plant MMR gene in vivo or to permit the sense molecule to participate in co-suppression. Alternatively, the MMR altering gene may encode a protein or proteins which interfere with the activity of a plant MMR protein and thus disrupt the plant's MMR system. For example, such encoded proteins may be antibodies or other proteins capable of interfering with MMR protein function, such as by 20 complexing with a protein functionally involved in plant MMR thereby disrupting the MMR of the plant. An example of such a protein is the MSH3 protein of Arabidopsis thaliana described herein or a protein of another plant which is homologous to the MSH3 protein of A. thaliana. For instance, overexpression of MSH3 in a plant cell causes MSH2 present in the cell to be substantially completely complexed, disrupting the 25 mismatch repair mechanism or mechanisms in the cell which are functionally dependent Similarly, mismatch repair on the presence of a complex of MSH2 with MSH6. mechanisms which depend on the presence of a complex of MSH2 and MSH3 may be disrupted by the overexpression of MSH6.

A chimeric gene of the fourth embodiment, incorporating a MMR altering gene, may be prepared by methods which are known in the art. Similarly, the MMR altering gene may be introduced into a plant cell, regenerating tissue or whole plant by techniques known in the art as being suitable for plant transformation, or by crossing. Known transformation techniques include Agrobacterium tumefaciens or A. rhizogenes mediated gene transfer, ballistic and chemical methods, and electroporation of protoplasts.

The MMR altering gene or genes are typically expressed from suitable promoters. Suitable promoters may direct constitutive expression, such as the 35S or the NOS promoter. Usually, however, the promoter will direct either inducible or tissue specific (e.g. callus; embryonic tissue; etc.), cell type specific (e.g. protoplasts; meiocytes; etc.) or developmental (e.g. embryo) expression of the altering gene or genes, in order for the

MMR system to function in tissue types or cell types, or at developmental stages of the plant, in which it is not desirable for the MMR system to be altered. Using such promoters, therefore, the activity of a MMR altering gene may be limited to a specific stage during plant development or it may be altered by controlling conditions external to the plant, and the deleterious effects of a permanently disabled or altered DNA mismatch repair system in a plant may be avoided. Examples of suitable promoters which are not constitutive are known in the art and include inducible promoters such as PR1a (reviewed by Gatz, 1997, Annual Rev. Plant Phys. Plant Mol. Biol. 48: 89), tissue specific promoters such as AoPR1 (Sabahattin et al., 1993, Biotechnology 11: 218), and cell-type specific promoters such as DMC1.

A chimeric gene in accordance with the invention may further be physically linked to one or more selection markers such as genes which confer phenotypic traits such as herbicide resistance, antibiotic resistance or disease resistance, or which confer some other recognisable trait such as male sterility, male fertility, grain size, colour, growth rate, flowering time, ripening time, etc.

The process of the tenth embodiment of the invention provides, for example, a process for generating intraspecies genetic variation by altering the mismatch repair system in a plant cell, in regenerating plant tissue or in a whole plant. The plant cell, regenerating tissue or whole plant includes and expresses one or more MMR altering 20 genes which are capable of altering mismatch repair in the plant cell, regenerating tissue or whole plant. Alteration of MMR may be achieved, for example, by inactivating the genes encoding plant MSH3 and/or plant MSH6. It is preferred to inactivate the plant MSH3 and MSH6 encoding genes at the same time and in the same plant cell, regenerating tissue or whole plant. Typically in this preferred form of the invention 25 inactivation of either plant MSH3 or MSH6 alone is insufficient to substantially alter the plant's mismatch repair system and only when both MSH3 and MSH6 are inactivated simultaneously is the plant's mismatch repair system sufficiently altered to prevent the MMR system from recognising base pair mismatches, base insertions or deletions as a result of DNA replication errors, DNA damage, or oligonucleotide induced site-specific 30 mutagenesis. However, in some applications of the invention, inactivation of only one gene may also be used to cause genomic instability or increase the efficiency of sitespecific mutagenesis.

If desired, the MMR altering gene or genes may later be rendered non-functional or ineffective, or may be removed from the genome of the plant cell, regenerating tissue or whole plant in order to restore mismatch repair in the plant cell, regenerating tissue or whole plant. The MMR altering gene or genes may be inactivated by means of known gene inactivation tools, such as ribozymes, or may be removed from the genome using gene elimination systems known in the art, such as CRE/LOX. It is preferred to render two genes, whose gene products combine to incapacitate MMR, ineffective by separating

the altering genes through segregation. Therefore, in a preferred embodiment of the invention a first plant cell or plant is generated in which only plant MSH3 is incapacitated, and a second plant cell or plant is generated in which only plant MSH6 is incapacitated. The combination of both genomes, for example by crossing, then produces significant 5 MMR deficiency in those cells or plants which have inherited both altering genes. If the altering genes are expressed from unlinked loci, gene segregation restores MMR activity in the progeny of the cells or plants.

In a process of the ninth embodiment of this invention, homeologous recombination is enhanced between different genomes, chromosomes or genes in plant cells or plants by altering MMR in said plant cells or plants. Such genomes, chromosomes or genes are characterised in that they originate from different plant families, genera, species, subspecies, plant varieties or lines. Hybrid plant cells or hybrid plants may be produced by crossing, by cell fusion or by other techniques known in the art. These plant cells or plants are further characterised by expressing one or more genes that are capable of altering mismatch repair in the plant cell or plants.

In the process of the ninth embodiment, the homeologous recombination is typically for the purpose of introducing a desired characteristic in the hybrid plant. In this typical application of the process of the ninth embodiment, and in the process of the tenth embodiment the desired characteristic may be any characteristic which is of value to the plant breeder. Examples of such characteristics are well known in the art and include altered composition or quality of leaf or seed derived storage products (e.g. oil, starch, protein), altered composition or quality of cell walls (e.g. decrease in lignin content), altered growth rate, prolonged flowering, increased plant yield or grain yield, altered plant morphology, resistence to pathogens, tolerance to or improved performance under environmental stresses of various kinds, etc.

In a preferred form of the tenth embodiment, an MMR altering gene is cointroduced along with the homeologous genome, chromosome or gene of another plant
cell or plant into an MMR proficient plant cell or MMR proficient plant to produce a
hybrid plant cell or hybrid plant in which homeologous recombination can occur.

Suitably, the MMR proficient plant cell or MMR proficient plant may also include an
MMR altering gene. For example a gene capable of inactivating plant MSH3 may be cointroduced along with the homeologous genome, chromosome or gene of another plant
cell or plant into an MMR proficient plant cell or MMR proficient plant in which MSH6
is inactivated. A resultant hybrid plant in which homeologous recombination occurs will
include both the MSH3 and MSH6 altering genes and its MMR system will therefore be
inactivated.

In this form of the invention, if hybrid plants are to be produced by crossing, the MMR altering gene preferably originates from the male parent, thus ensuring that the

MMR altering gene is always introduced and is not present in the recipient cell. That is, the MMR of the recipient cell, prior to introduction of the MMR altering gene, is typically proficient. Alternatively, if an MMR altering gene is present in a recipient cell it may be ineffective or inefficient on its own, or it may be linked to an inducible or tissue specific or cell type specific promoter which only renders the MMR altering gene active under limited conditions.

Thus, in a preferred form of the process of the ninth embodiment, the MMR system of the hybrid plant is initially unaltered. In this form of the process, the step of altering the mismatch repair system may comprise introducing into the hybrid plant, or cells thereof, a MMR altering gene, such as by Agrobacterium tumefaciens or A. rhizogenes mediated gene transfer, ballistic and chemical methods, and electroporation of protoplasts.

The MMR altering gene or genes are typically expressed from suitable promoters, as described above. Preferably, the promoter is transcriptionally active in mitotically and meiotically active tissue and/or cells to ensure MMR alteration after chromosome pairing at mitosis and meiosis, respectively. The preferred timing for MMR alteration is at meiosis, because recombinant genomes, chromosomes or genes are directly transmitted to the progeny. A suitable meiocyte specific promoter is for example the DMC1 promoter from Arabidopsis thaliana ssp. Ler. (Klimyuk and Jones, 1997, Plant J. 11, 1-14). However, mitotic homeologous recombination is also a desirable outcome as somatic recombination events can be transmitted to offspring due to the totipotency of plant cells and the lack of predetermined germ cells in plants.

If desired, the MMR altering gene or genes may later be rendered non-functional or ineffective, or may be removed from the hybrid plant or hybrid plant cells, in order to restore mismatch repair in the hybrid plant or hybrid plant cells. The MMR altering gene or genes may be inactivated by means of known gene inactivation tools as described herein above.

EXAMPLES

Example 1. Cloning of the AtMSH3 and AtMSH6 coding sequences

Isolation of partial AtMSH3 and AtMSH6 consensus sequences

Degenerate oligonucleotides UPMU (SEQ ID NO:1) and DOMU (SEQ ID NO:2)

UPMU CTGGATCCACIGGICCIAA(C/T)ATG

DOMU CTGGATCC(A/G)TA(A/G)TGIGTI(A/G)C(A/G)AA

were used to isolate AtMSH3 and AtMSH6 sequences by PCR amplification.

Primers UPMU and DOMU correspond to conserved amino acid sequences of the proteins MutS (E. coli and S. ryphimurium), HexA (S. pneumoniae). Repl (mouse) and Ducl (human). The conserved regions to which they are targeted are TGPNM for UPMU (amino acid positions 852-856 for AtMSH6 and 816-820 for AtMSH3) FATHY or FVTHY

for DOMU (amino acid positions 964-968 for AtMSH6 and 928-932 for AtMSH3, respectively.) These primers have been used to isolate MSH2 and MSH1 from yeast (Reenan and Kolodner, Genetics 132: 963-973 (1992)) and MSH2 from *Xenopus* and mouse (Varlet et al., Nuc. Acids Res. 22:5723-5728 (1994)).

Template single strand cDNA was produced by reverse transcription of 2 µg total RNA from a cell suspension culture of Arabidopsis thaliana ecotype Columbia (Axelos et al. 1989, Mol. Gen. Genetics 219: 106-112). The PCR reaction was performed under the following conditions in a final volume of 100µl: 0.2mM dNTP, 1µM each primer, 1XPCR buffer, lu Taq DNA polymerase (Appligene) in the presence of template cDNA. PCR 10 parameters were 5 minutes at 94°C, followed by 30 cycles of 40 seconds at 95°C, 90 seconds at 45°C, 1 minute at 72°C. The amplification products were cloned into pGEM-T vector (Promega) and sequenced. Two different clones were isolated, S5 (350bp) was homologous to MSH3, S8 (327bp) was homologous to MSH6. Complete cDNA sequences were then isolated according to the Marathon cDNA amplification kit procedure (Clontech). 15 In summary, this procedure involves producing double stranded cDNA by reverse transcription of 2µg polyA+ RNA from the cell suspension culture of Arabidopsis. Adaptors are ligated on each side of the cDNA. The ligated cDNA is used as a template for 5' and 3' RACE PCR reactions in the presence of primers that are specific for the adaptor on one side (AP1 and AP2), and specific for the targeted gene on the other side. A 5' and a 3' 20 fragment that overlap are thus produced for each gene. The complete gene coding sequence can be reconstituted taking advantage of a unique restriction site, if available, in the overlapping region. Specific details of this procedure as it was used to isolate AtMSH3 and AIMSH6 coding regions, are as follows.

Isolation of AtMSH3 complete coding sequence

From the sequence of clone S5, primer 636 (SEQ ID NO:3) was designed:

636 TGCTAGTGCCTCTTGCAAGCTCAT.

Primer AP1 (SEQ ID NO:4) is complementary to a portion of an adaptor sequence which had been ligated to the 5' and 3' ends of Arabidopsis cDNA:

API CCATCCTAATACGACTCACTATAGGGC.

PCR performed on the ligated cDNA with primers 636 and AP1 for the 5' RACE PCR was followed by a second round of amplification with the nested primers AP2 (SEQ ID NO:5) and S525 (SEQ ID NO:6)

AP2 ACTCACTATAGGGCTCGAGCGGC

S525 AGGTTCTGATTATGTGTGACGCTTTACTTA

35 (the latter was also designed to correspond to a part of the sequence of clone S5) and produced a 2720bp DNA fragment. Figure 1 provides a diagrammatic representation of the primer sequences used to isolate AtMSH3. Another primer (S51, SEQ ID NO:7)

S51 GGATCGGGTACTGGGTTTTGAGTGTGAGG

25

was designed closer to the 5° border and permitted the determination of 99bp upstream to the ATG initiation codon. For the 3° RACE PCR, a first PCR reaction was performed with primers AP1 and 635 (SEQ ID NO:8).

- 635 GCACGTGCTTGATGGTGTTTTCAC
- 5 followed by a second round of amplification, using the nested primers AP2 and S523 (SEQ ID NO:9)
- which produced a DNA fragment of 890bp. Both DNA fragments were subcloned into pGEM-T and sequenced. Since PCR amplification using the Expand Long Template PCR system (Boehringer-Mannheim) produced errors in the sequence, new oligonucleotides were designed to isolate those sequences again by PCR, but with the high fidelity DNA polymerase Pfu. PCR with primers 1S5 (SEQ ID NO:10) and S53 (SEQ ID NO:11)
 - 1S5 ATCCCGGGATGGGCAAGCAAAAGCAGCAGACGA
 - S53 GACAAAGAGCGAAATGAGGCCCCTTGG
- amplified the 1244bp fragment clone 52 (SEQ ID NO:12, cloned into pUC18/Sma1). PCR with primers S52 (SEQ ID NO:13) and 2S5 (SEQ ID NO:14)
 - 2S5 ATCCCGGGTCAAAATGAACAAGTTGGTTTTAGTC
 - S52 GCCACATCTGACTGTTCAAGCCCTCGC

amplified the 2104bp clone 13 (SEQ ID NO:15, cloned into pUC18/Sma1). The complete coding sequence of the AtMSH3 gene was reconstructed in pUC18 by ligating the 5' half of AtMSH3 (clone 52) to the 3' half of AtMSH3 (clone 13) after digesting with BamH1 which has a unique cleavage site in the overlapping region of both clones. This manipulation yielded plasmid pPF26. The Smal fragment from pPF26 contains the complete AtMSH3 coding sequence. The remaining primers referred to in Figure 1 are as follows:

- S51 GGATCGGGTACTGGGTTTTGAGTGTGAGG (SEQ ID NO:16)
- S525 AGGTTCTGATTATGTGTGACGCTTTACTTA (SEQ ID NO:17)

Figures 2 and 3 provide plasmid maps of clones 52 and 13 respectively, showing restriction enzyme cleavage sites. The complete AtMSH3 coding sequence (SEQ ID NO:18) 30 is 3246bp long and is shown in Figure 4 together with the deduced sequence (SEQ ID NO:19) of the encoded polypeptide. AtMSH3 is clearly homologous to the yeast and mouse MSH3 genes. A sequence alignment of polypeptides encoded by AtMSH3 and that encoded by Saccharomyces cerevisiae MSH3 is set out in Figure 5.

Isolation of the AtMSH6 complete coding sequence and genomic sequences

- The same procedure allowed isolation of the AtMSH6 cDNA. Figure 6 provides a diagrammatic representation of the primer sequences used to isolate AtMSH6. For the 5 RACE PCR, primers 638 (SEQ ID NO:20) and AP1 (SEQ ID NO:4)
 - 638 TCTCTACCAGGTGACGAAAAACCG
 allowed the amplification of a 2889 DNA fragment. Primer S81 (SEQ ID NO:21)

S81 CGTCGCCTTTAGCATCCCCTTCCTTCAC

helped define the 142bp upstream to the ATG initiation codon. On the 3' side, RACE PCR was initially performed with primers S823 (SEQ ID NO:22) and AP1 (SEQ ID NO:4).

S823 GCTTGGCGCATCTAATAGAATCATGACAGG

5 and then with the nested primers 637 (SEQ ID NO:23) and AP2 (SEQ ID NO:5).

637 GACAGCGTCAGTTCTTCAGAATGC

to produce a 774bp DNA fragment. As for AtMSH3, those fragments were cloned and sequenced. Re-isolation of the DNA sequence using the high fidelity Pfu polymerase and newly designed primers 1S8 (SEQ ID NO:24) and S83 (SEQ ID NO:25) (for the 5° side) led to a 2182 bp DNA fragment identified as clone 43 (SEQ ID NO:26, cloned in pUC18/Sma1), and a 1379bp clone identified as clone 62 (SEQ ID NO:27, also cloned in pUC18/Sma1).

1S8 ATCCCGGGATGCAGCGCCAGAGATCGATTTTGT

2S8 ATCCCGGGTTATTTGGGAACACAGTAAGAGGATT (SEQ ID

NO:28)

15

S82 GCGTTCGATCATCAGCCTCTGTGTTGC (SEQ ID NO:29)

S83 CGCTATCTATGGCTGCTTCGAATTGAG

Figures 7 and 8 provide plasmid maps of clones 43 and 62 respectively, showing restriction enzyme cleavage sites. Clones 43 and 62 were digested by the Xmn1 restriction enzyme for which a unique site is present in their overlapping region and then ligated. The complete AtMSH6 coding sequence (SEQ ID NO:30) is 3330bp long and is shown in Figure 9 together with the deduced sequence (SEQ ID NO:31) of the encoded polypeptide. AtMSH6 is clearly homologous to the yeast and mouse MSH6genes. A sequence alignment of polypeptides encoded by AtMSH6 and that encoded by Saccharomyces cerevisiae MSH6 is set out in Figure 10.

An AtMSH6 genomic sequence was also isolated from a genomic DNA library constituted after partial Sau3Al digestion of DNA from the Arabidopsis cell suspension. 8062bp were sequenced that covered the AtMSH6 gene and show colinearity with the cDNA. 16 introns are found scattered along the gene. The complete genomic sequence 30 (SEQ ID NO:98) is shown in Figure 11.

Example 2. A measure of somatic variation in MMR deficient plants

Constructs

Constructs with antisense AtMSH3 or antisense AtMSH6 or both AtMSH3/AtMSH6 under the control of a single 35S promoter have been inserted into the binary vector pPZP121 (Hajdukiewicz et al., Plant Mol. Biol. 23, 793-799) between the right and left borders of the T-DNA. The pPZP121 plasmid confers chloramphenicol resistance to Escherichia coli or Agrobacterium tumefaciens bacteria. The aacC1 gene is carried by the T-DNA and allows selection of transformed plant cells on gentamycin (Hajdukiewicz et al., Plant Mol. Biol. 25, 989-994). For the purpose of expressing antisense constructs, a 35S

promoter/terminator cassette with a polylinker was introduced into pPZP121. The 3' ends of the considered genes have been chosen since this region seems more efficient for antisense inhibition. For AtMSH3 this corresponds to clone 13 (2104bp), for AtMSH6 this corresponds to clone 62 (1379bp). Clone 13 comprises 2104bp of the 3' region that were cut off the pUC18 vector by Sal1/Sst1 restriction, blunted with T4 DNA polymerase and ligated into the T4 DNA polymerase blunted BamHI site of pPZP121/35S, creating clone pPF13. The same procedure was followed for the 3' region of AtMSH6 clone 62 (1379bp) thus creating plasmid pPF14. For the double constructs, the 3' region (from clone 62) of AtMSH6 was introduced ahead of the AtMSH3 region into pPF13 creating pCW186 and reciprocally, the 3' region of AtMSH3 (from clone 13) was introduced ahead of AtMSH6 into pPF14, creating pCW187.

These constructs were introduced into the Arabidopsis cells (as described below) of wildtype Columbia and of the Columbia tester line.

An alternative strategy to antisense inhibition of AtMSH6 comes from experiments of 15 Marra et al. (1998, Proc. Natl. Acad. Sci USA 95, 8568-8573) who show that overexpression of functional MSH3 results in depletion of MSH6 protein in human cells. This depletion may generate a mismatch repair mutant phenotype.

For the purpose of overexpressing functional AtMSH3 protein in plant cells, the complete MSH3 coding region was excised from pPF26 (example 1) by digestion with 20 Smal, and was inserted into the Smal site of pCW164. The resulting construct was named pPF66. It contains a complete AtMSH3 gene under the control of the 35S promoter inside the left (LB) and right (RB) border of the T-DNA. This T-DNA also contains the hpt2 gene for gentamycin selection. Plasmid pPF66 was introduced into Arabidopsis cells as described below. One cell clone was selected which clearly overexpressed the AtMSH3 gene as shown by Northern analysis. Figures 12-16 provide plasmid maps of plasmids pPF13, pPF14, pCW186, pCW187 and pPF66. respectively.

Construction of tester construct

For the purpose of Forward Mutagenesis Assays, a tester construct was built containing the coding regions for nptII, codA, uidA. All three genes are driven by the 35S promoter and are terminated by the 35S terminator. This construct was obtained by introducing an EcoR1 fragment encoding the codA cassette (2.5kb) and a HindIII fragment encoding the uidA (GUS) cassette (2.4kb) into the pPZP111 vector (Hajdukiewicz et al.,1994, Plant Mol Biol 23: 793-799) which already contained the nptII expression cassette. This new plasmid was named pPF57. NptII is used to select for transformed plant cells.

35 GUS is used to analyse the degree of gene silencing in the construct (i.e. to identify cell lines in which the transgenes are expressed), and codA is used as a marker for forward mutagenesis (described below).

The plasmid map of pPF57 is provided in Figure 17.

Plant cell transformation

then transformed by co-cultivation. A suspension culture of *Arabidopsis thaliana* cells that 5 has been established by Axelos et al. (1992, Plant Physiol. Biochem. 30, 1-6) may be used. One day old freshly subcultured cells are diluted five times into AT medium (Gamborg B5 medium. 30g/l sucrose, 200µg/l NAA). 10µl of saturated *Agrobacterium* containing the transforming T-DNA constructs are added to 10ml diluted cells in a 100ml erlenmeyer. The co-cultivation is agitated slowly (80rpm) for 2 days. The cells are then washed 3 times into AT medium and finally resuspended in the same initial volume (10ml). The culture is agitated for 3 days to allow expression before plating on selection plates (AT/BactoAgar 8g/l+gentamycin 50µg/ml). Transformed individual calli are isolated 3 weeks later.

Tester Strain

The tester construct on plasmid pPF57 was introduced into Arabidopsis cells of wildtype Columbia using the transformation protocol described above. Among 10 candidate transformants, one cell clone was shown (by Southern analysis) to have a unique T-DNA insertion. All three genes were shown to be functional in this cell line as indicated by resistance to kanamycin, blue staining in the presence of X-Glu (GUS), and sensitivity to 5-fluoro-cytosine (codA).

MMR altering genes (described above) were then introduced individually into the tester line and transformed cells are used for analysis of both Microsatellite Instability and Forward Mutagenesis.

Microsatellite analysis

Microsatellites have been described in Arabidopsis (Bell and Ecker, 1994, Genomics 19, 137-144). The present Example is based on a study of instability of microsatellites that are polymorphic amongst different ecotypes. DNA is extracted from the transformed calli. Specific primers have been defined that are used to amplify the microsatellite sequence. One of the two primers is previously P³² labelled by T4 kinase. In case of a polymorphic variation, new PCR products appear that do not follow the expected pattern of migration on a polyacrylamide gel. This is a commonly observed feature for MMR deficient cells in yeast or mammalian cells.

In particular, the present Example describes a study on microsatellites ca72 (CA₁₈), nga172 (GA₂₉), and ATHGENEA(A₃₉), chosen because they belong to the types predominantly affected in human mismatch repair deficient tumors. The size of these microsatellites is not conserved from one *Arabidopsis* ecotype to the other.

Arabidopsis cells which are transformed with an MMR altering gene (above) and control cells not expressing the MMR altering gene are allowed to form calli. DNA is

rapidly extracted from the calli and is analysed for microsatellite instability as described in detail by Bell and Ecker 1994. Genomics 19, 137-144. In summary, the relevant microsatellite is amplified by PCR using P32 labelled primers. The PCR products are separated on a DNA sequencing gel for size determination. Size differences between microsatellites from transformed and control cells not expressing the MMR altering gene in question indicate microsatellite instability as a result of MMR alteration.

The sequences of primers used for PCR amplification of microsatellites ca72 and nga172 are included in Table 1. PCR amplification of microsatellite ATHGENEA made use of a forward primer containing the sequence

ACCATGCATAGCTTAAACTTCTTG (SEQ ID NO:32)

and of a reverse primer containing the sequence

ACATAACCACAAATAGGGGTGC (SEQ ID NO:33).

The amplification for microsatellite ca72 revealed in Columbia control cells (with respect to the MMR altering gene) a 248 bp long PCR fragment instead of the published length of 124 bp. DNA sequencing verified this fragment as a CA₁₈ microsatellite.

Forward mutagenesis assay

10

Tester cells transformed with antisense AtMSH3 or antisense AtMSH6 or both AtMSH3/AtMSH6 are analysed for the stability of the codA gene. The functional codA gene confers to sensitivity to 5-fluoro-cytosine (5FC), whereas a gene inactivating mutation in codA will confer resistance to 5FC. The frequency of resistant cells is therefore a good indicator of somatic variation as a direct result of MMR alteration. Variants resistant to 5FC are first analysed for GUS activity. If GUS is inactive, 5FC resistance is assumed to be due to gene silencing (all three genes are under the 35S promoter). If GUS is active, 5FC resistance is assumed to be due to forward mutations that have inactivated codA. PCR is then performed on the putative codA mutant genes which is then sequenced to confirm the presence of forward mutations in codA.

Besides codA, other marker genes may also be used for the Forward Mutagenesis Assay such as the ALS gene (conferring sensitivity to valine or to sulfonylurea; Hervieu and Vaucheret, 1996, Mol. Gen. Genet. 251 220-224; Mazur et al. 1987, Plant Physiol. 85 1110-30 1117).

Example 3. Homeologous meiotic recombination in Arabidopsis thaliana

A. Construction of a meiocyte specific gene expression cassette comprising the DMC1 promoter and the NOS terminator

(i) The *DMC*1 promoter may be used as published by Klimyuk and Jones, 1997, 35 Plant J. 11.1-14). To obtain a more convenient alternative for gene cloning, a 3.3 Kb

long subfragment of the DMC1 promoter was obtained by PCR from genomic DNA of Arabidopsis thaliana (ssp. Landsberg erecta "Ler").

The PCR was done in three rounds:

Round One: A 3.7 Kb long product was obtained using the forward primer 5 DMCIN-A comprising the sequence

GAAGCGATATTGTTCGTG (SEQ ID NO:34)

and the reverse primer DMCIN-B comprising the sequence

AGATTGCGAGAACATTCC (SEQ ID NO:35).

The weak amplification product was then used as template for round two and three.

Round Two: A 3.1 Kb long product comprising the promoter and the 5' untranslated leader was obtained using forward primer DMCIN-1, which contained the sequence

acgcgtcgacTCAGCTATGAGATTACTCGTG (SEQ ID NO:36)

and introduced a SaII cloning site at the 5' end of the promoter fragment, and reverse primer DMCIN-2 which contained the sequence

gctctagaTTTCTCGCTCTAAGACTCTCT (SEQ ID NO:37)

and introduced a XbaI site at the 3' end of the PCR fragment.

Round Three: A 0.2 Kb long product comprising the first exon/intron of the DMC1 promoter was obtained using forward primer DMCIN-3, which contained the sequence

gctctagaGCTTCTCTTAAGTAAGTGATTGAT (SEQ ID NO:38)

and introduced a XbaI site at the 5' end of the PCR fragment, and reverse primer DMCIN-4, containing the sequence

tcccccgggctcgagagatctccatggTTTCTTCAGCTCTATGAATCC (SEQ ID NO:39) and introduced at the 3' end of the PCR product restriction sites for Ncol, Bg/II. Xhol and 25 Smal.

The products obtained in round Two and Three were digested with XbaI and subsequently ligated to reconstitute a 3.3 Kb long DMC1 promoter from which the first two in-frame ATG start codons were replaced with a unique restriction site for XbaI. This promoter can be cloned between the restriction sites for SalI and SmaI of p3264, which contains the SacI-EcoRI NOS terminator in pBIN19, to yield the entire expression cassette in pBIN19. This cassette contains the following cloning sites: NcoI, BglII, XhoI. SmaI and (already present on p3264) KpnI and SacI.

(ii) Another strategy yielded the following convenient *DMC*1 promoter. A 1.8 kb DNA fragment comprising the 3' terminal part of the meiocyte specific *DMC*1 promoter was isolated by PCR from purified genomic DNA of *Arabidopsis thaliana* (ssp. Landsberg erecta "Ler"). The forward PCR primer (DMC1a) contained the sequence

acgcgtcgacGAATTCGCAAGTGGGG (SEQ ID NO:40)

and introduced a Sall cloning site at the 5' end of the promoter fragment. The reverse PCR primer (DMC1b) contained the sequence

20

tecatggagatetecegggtacCGATTTGCTTCGAGGG (SEQ ID NO:41)

introducing a polylinker region at the 3' end of the promoter fragment. The PCR reaction was carried out with VENT DNA Polymerase (NEB) over 25 cycles using the following cycling protocol: 1 minute at 94°C, 1 minute at 54°C, 2 minutes at 72°C.

The PCR reaction yielded a blunt ended DNA fragment which was digested with restriction enzyme SaII and was cloned into the cleavage sites of restriction enzymes SaII and SmaI in plasmid p2030, a pUC118 derivative containing the SacI-EcoRI NOS terminator fragment of pBIN121. The cloning yielded plasmid p2031, containing the DMC1 promoter-polylinker-NOS terminator expression cassette depicted in Figure 18.

10 B. Construction of an MSH3 antisense gene under the control of the DMC1 promoter

A 2.1 kb DNA fragment encoding the carboxyterminal part of AtMSH3 was removed from the polylinker of clone 13 described in Example 1 after (i) digestion with KpnI. (ii) blunting of the DNA ends generated by KpnI and (iii) digestion with BamHI. The isolated fragment was then cloned in antisense orientation downstream of the DMC1 promoter in plasmid p2031, which had been digested with restriction enzymes SmaI and BglII. This cloning yielded plasmid p2033 (Figure 18).

After digestion of p2033 with EcoRI, a 4.1 kb DNA fragment was recovered comprising the DMC1 promoter, the partial MSH3 cDNA in antisense orientation with respect to the promoter and the NOS terminator. This fragment was cloned into the EcoRI restriction site of plant transformation vector pNOS-Hyg-SCV to yield plasmid p3242 (Figure 18).

C. Construction of a combined MSH6/MSH3 antisense gene under the control of a single DMC1 promoter

A 3.1 kb fragment, encoding in antisense orientation the partial AtMSH6 and AtMSH3 sequences and the 35S terminator, was isolated from pCW186 by digestion with KpnI. The fragment was treated with Klenow enzyme to blunt both ends. It was then cloned into the SmaI site of plasmid p3243 (a pNOS-Hyg-SCV derivative, illustrated in Figure 19), which had been opened to delete the region between the SmaI sites. Clones containing the fragment in the antisense orientation with respect to the DMC1 promoter (described in A(ii) above) were identified by diagnostic digestion with BamHI. The selected construct was named p3261.

Another practical way of cloning the double antisense gene is as follows. A 1 kb DNA fragment encoding the carboxyterminal part of AtMSH6 is isolated from clone 62 described in Example 1 after digestion of clone 62 plasmid DNA with BamHI, which cleaves in the 5' polylinker region flanking the partial cDNA, and with EcoRI, which cleaves within the cDNA. The isolated fragment is treated with Klenow enzyme to blunt both its ends and is cloned into the recipient plasmid p2033 or p3242. For the purpose of

cloning, the recipient plasmid may be cleaved with either Aval or Ncol and can be blunted with Klenow enzyme to produce blunt acceptor ends for fragment cloning. This cloning yields two opposite orientations of cloned fragment DNA with respect to the DMC1 promoter. These can be identified by diagnostic digestion with restriction enzymes Scal or Xmnl in conjunction with Sacl. The selected construct contains the DMC1 promoter, the combined partial cDNAs for AtMSH3 and AtMSH6 (both cloned in antisense orientation with respect to the DMC1 promoter) and the NOS terminator. If the recipient plasmid is p2033, the combined antisense gene under control the single DMC1 promoter is recovered from the vector after EcoRI digestion and is cloned into the EcoRI restriction site of pNOS-Hyg-SCV.

D. Construction of a full-length MSH3 sense gene under control of the DMC1 promoter for overexpression of functional MSH3 protein

Overexpression of MSH3 protein was shown in human cells (Marra et al., 1998, Proc. Natl. Acad. Sci. USA 95, 8568-8573) to complex all available MSH2 protein. This leaves MSH6 protein without its partner, leading to the degradation of MSH6 protein and eventually to a mismatch repair phenotype.

This phenomenon is exploited to increase homeologous meiotic recombination in Arabidopsis as an alternative to antisense inhibition of AtMSH6. For this purpose the full-length cDNA encoding AtMSH3 is isolated from plasmid pPF66 by digestion with Smal and is cloned into the Smal site of the DMC1 expression cassettes described in A(i).

E. Selection of Recombination markers on homeologous chromosomes of Arabidopsis thaliana subspecies Landsberg erecta (Ler), Columbia (Col) and C24, respectively

E(i). Visual recombination markers in Arabidopsis th. subspecies C24:

Agrobacterium mediated transformation with a T-DNA containing a 35S-GUS gene, inactivated by insertion of a 35S-Ac transposable element (Finnegan et al., 1993, Plant Mol. Biol. 22, 625-633), had yielded a C24 line in which the T-DNA construct was integrated into chromosome 2. Genetic and molecular analysis of this line shows that the Ac transposon had excised from its T-DNA locus thereby restoring GUS activity, but had re-inserted into the chromosome at a distance of 16.4 cM, where it stayed fixed (due to disablement of Ac) within the chlorina gene. Insertional inactivation of the chlorina gene caused a bleached phenotype in those plants that were homozygous for this mutation. Because of the two linked phenotypic markers, chlorina3A:Ac and GUS, this C24 line was used in crosses to wildtype Ler for analysis of meiotic homeologous recombination on chromosome 2 in conjunction with molecular recombination markers.

35 E(ii) Visual recombination markers in Arabidopsis th. Ler:

The Ler line NW1 (obtained from NASC, Nottingham, UK) contains one recessive visual marker per chromosome. i.e. an-1 on Chr.1, py-1 on Chr.2, gl1-1 on Chr.3, cer2-1

10

on Chr.4, and ms1-1 on Chr.5. This line is used in crosses to wildtype C24 which expresses an MMR altering gene for analysis of meiotic homeologous recombination on chromosomes 1-5 in conjunction with molecular recombination markers listed in Table 1.

Other Ler lines from NASC have several visual markers in close proximity to each other on the same chromosome. When these lines are used for hybrid production, analysis of homeologous meiotic recombination may make use entirely of visual recombination markers. Particularly suitable for crossing to C24 wildtype that is expressing a MMR altering gene are the following Ler lines:

NW22: relative markers are dis1 - (4 cM) - ga4 - (11 cM) - th1 on chromosome 1.

NW10: relevant markers are tz-201 - (5 cM) - cer3 on chromosome 5.

NW134, relevant markers are ttg - (4 cM) - ga3 on chromosome 5.

NW24 (abi3-1) and NW64 (gl1-1). When present in the same plant on chromosome 3. abi3-1 and gl1-1 are 13 cM apart. Since this marker combination is not available from NASC, we have combined these markers by crossing line NW24 to line NW64. The F1 offspring were allowed to self-fertilise and to produce F2 seeds. F2 Plants which carry both markers as homozygous traits on the same chromosome can be identified firstly, by germinating F2 seeds on germination medium containing selective concentrations of abscisic acid, and subsequently, by identifying among the abscisic acid resistant plants those individuals which show the glabra phenotype.

20 E(iii) Molecular recombination markers in Col, Ler and C24:

The genome of Arabidopsis thaliana is interspersed with unique base sequences arranged as simple tandem repeats. Allelic repeats can vary in length between different Arabidopsis subspecies and when amplified by PCR yield diagnostic DNA products of different length named Simple Sequence Length Polymorphisms (SSLPs). Many SSLPs have been genetically mapped and have been assigned to unique chromosome locations on the recombinant inbred map (Bell and Ecker, 1994, Genomics 19, 137-144; Lister and Deans lines, Weeds World 4i, May 1997).

In Table 1 are listed 28 mapped and established SSLPs between Ler and Col. A number of PCR primer pairs are described herein (SEQ ID NO:42 to SEQ ID NO:97) which also yielded SSLPs between C24 and Ler (19 SSLPs) or between C24 and Col (25 SSLPs), respectively. Polymorphic SSLPs can be used as molecular markers in the analysis of homeologous recombination between genomes from these subspecies.

The PCR reactions (25 μL) were carried out over 35 cycles (15 seconds at 94°C, 30 seconds at 55°C and 30 seconds at 72°C), with 0.25 U Taq DNA polymerase and 0.6 μg genomic DNA in reaction buffer containing 2 mM MgCl₂. PCR products were separated by agarose gel electrophoresis (4% ultra high resolution agarose) and visualised by ethidiumbromide staining. The results from the PCR experiments are summarised in

Table 1, which also shows the sequence of PCR primers, primer annealing temperature (Tm). PCR product length and chromosome location of SSLP (with respect to the RI map of May 1997, Weeds World 4i).

F. Production of hybrid plants

C24 plants heterozygous for chlorina3A:Ac/GUS are crossed as male to emasculated wildtype Ler to produce Ler/C24(chlorina3A, GUS) hybrid seeds.

Due to the heterozygosity of the C24 parent, only 50 % of hybrid plants have inherited the chlorina3A:Ac/GUS locus. The remaining 50% of hybrid plants are wildtype with respect to chlorina3A:Ac/GUS. Since the mutant locus is linked to a kanamycin resistance gene (contained on the same T-DNA as GUS) mutant plants can be pre-selected by germinating hybrid seeds on germination medium containing 50 mg/L kanamycin.

Ler plants homozygous for the five chromosome markers are male sterile (ms1-1) and are crossed without emasculation to wildtype C24 to produce Ler(an-1, py-1, gl1-1, cer2-1, ms1-1)/C24 hybrid seeds.

Other Ler plants, which are male fertile, are crossed after emasculation of the female parent to produce Ler/C24 hybrid seeds.

G. Introduction of MSH3 and MSH6/3 antisense genes into Arabidopsis and analysis of meiotic homeologous recombination

(i) Transformation of hybrid plants and analysis of homeologous meiotic recombination

The plant transformation vectors comprising the antisense genes described in (B) and (C) above are introduced into Agrobacterium tumefaciens strain AGL1 (Lazo et al., 1991, Bio/Technology 9, 963-967) by electroporation. Recombinant Agrobacterium clones are selected on LB medium containing 50 mg/L rifampicin and 100 mg/L carbenicillin. Selected clones are used to infect roots of Arabidopsis hybrid plants (described in (F) above) using the root transformation protocol of Valvekens et al. (1988, PNAS 85, 5536-5540) except that shoot and root inducing media contain hygromycin (10 mg/L) instead of kanamycin.

Plants regenerated from roots of hybrid plants are genetic clones of root donating plants and therefore are again genetic hybrids of two Arabidopsis subspecies described in 30 (F). However, in contrast to the root donating plants, the regenerated hybrid plants also contain the introduced transgene and the co-introduced hygromycin resistance gene with the latter allowing these plants to grow on hygromycin containing culture medium.

Hygromycin resistant plants are then allowed to enter the reproductive phase and to produce gametes by meiotic divisions of microspore and megaspore mothercells. At meiosis, the *DMC*1 promoter is activated and can direct the expression of antisense genes described in (B) and (C) above, leading to decreased *MSH*6 and/or *MSH*3 gene

expression. This in turn depletes the gamete mothercells of MSH6 and/or MSH3 protein, thus causing alteration of MMR during meiotic divisions and increasing the recombination frequency between homeologous chromosomes.

Transgenic plants are then allowed to self-fertilise and to produce seeds. These seeds (F2 seeds with respect to hybrid production), and the plants derived therefrom, carry the homeologous recombination events which can be identified by using the visual and molecular recombination markers described in (E) above.

In case of homeologous recombination between chromosomes of Ler and C24(chlorina3A:Ac, GUS), the analysis concentrates on chromosome 2 by selecting plants showing the visual phenotypic marker chlorina. This marker thus serves as a reference point as it indicates that respective chromosomes 2 originate from C24. Other markers, such as GUS or molecular markers, on chromosome 2 may then be used to identify chromosomal regions which are derived from the Ler chromosome as a result of homeologous recombination. F2 plants of control transformants not expressing the antisense gene(s) can be analysed in parallel and the results can be used for comparison to homeologous recombination results obtained in antisense plants.

(ii) <u>Transformation of C24 wildtype</u>, hybrid plant production and analysis of homeologous meiotic recombination

Introduction of MMR altering genes into wildtype C24 is done using the root transformation protocol as described in G(i) for transformation of hybrid plants. Transformed plants are selected by resistance to either 10 mg/L hygromycin (in case of transformation with T-DNA's derived from pNOS-Hyg-SCV) or to 50 mg/L kanamycin (in case of transformation with T-DNA's derived from pBIN19).

Transgenic plants are then allowed to self-fertilise and to produce seeds (T1 seeds).

Segregation of the antibiotic resistance gene in the T1 population then indicates the number of transgene loci. Lines with a single transgene locus (indicated by a 3:1 ratio of resistant:sensitive plants) are selected and are bred to homozygosity. This is done by collecting selfed seeds (T2) from T1 plants and subsequent testing of at least four independent T2 seed populations for segregation of the antibiotic resistance gene. T2 populations which do not segregate the antibiotic resistance gene are assumed to be homozygous for both the resistance gene and the linked MMR altering gene.

C24 plants homozygous for the MMR altering gene are then crossed to Ler lines homozygous for recessive visual markers (see E(ii)) to produce C24/Ler hybrid plants as described in (F). These F1 hybrids are then allowed to enter the reproductive phase and to produce gametes by meiotic division of microspore and megaspore mothercells. At meiosis, the DMC 1 promoter is activated and can direct the expression of antisense or sense genes described in (B), (C) and (D) above, leading to decreased MSH6 and/or MSH3 gene expression. This in turn depletes the gamete mothercells of MSH6 and/or MSH3

protein, thus causing alteration of MMR during meiotic divisions and increasing the recombination frequency between the homeologous chromosomes of C24 and Ler. Recombination events are then scored in the F2 generation.

For recombination analysis, the hybrid plants are allowed to self-fertilise and to produce F2 seeds. F2 plants are then preselected for a first visual marker. Since this marker is recessive, its visual presence indicates homozygosity for Ler DNA at the relevant locus. Those F2 plants which show this first visual marker are then analysed for the presence or absence of a second visual marker which in the Ler parent is closely linked to the first marker. Absence of the second visual marker indicates recombination between the relevant C24 and Ler chromosomes between the first and second marker. The frequency of recombination in transgenic hybrids is compared to the recombination frequency in control hybrids not expressing the MMR altering gene.

Examples of recombination analysis are the following.

The Ler line NW22(dis1. ga4. th1) is used for crosses to transformedC24.

F2 plants are preselected first for thiamine requirement (th1) and then are further analysed for re-appearance of wildtype height (loss of ga4) and/or re-appearance of normal trichomes (loss of dis1) as a result of recombination.

The Ler line NW10(1z-201, cer3) is used for crosses to transformed C24.

F2 plants are then preselected first for thiazole requirement (tz) and then are further 20 analysed for re-appearance of normal. i.e. non-shiny stems (loss of cer3) as a result of recombination.

The Ler line NW134 (ttg, ga3) is used for crosses to transformed C24. F2 plants are first preselected for dwarfish appearance (ga3) and are then analysed for re-appearance of trichomes (loss of ttg) as a result of recombination.

Ler plants homozygous for abi3-1 and gl1-1 are used for crosses to transformed C24. F2 plants are first preselected for their ability to germinate in the presence of abscisic acid and are then analysed for re-appearance of trichomes on the leaves (loss of gl1-1) as a result of recombination.

In the case of homeologous recombination between transformed C24 and the Ler line 30 NW1 (an-1, py-1, gll-1, cer2-1, msl-1), recombination analysis is similar the one described above, except that the second marker is not a visual marker but has to be a molecular marker. This is because the Ler parent carries only one visual marker per chromosome.

		ABLE 1: SS	TABLE 1: SSLP Markers in Arabidopsis thaliana Subspecies	haliana S	ubspecies		
Chromosome	RI Map Position	PCR Primer Pair	Primer Sequence	Tm (°C)	length/COL	length/LER	length/C24
_	2.3	ATEAT! F	GCCACTGCGTGAATGATATG	57.8	172	162	162
		ATEAT! R	CGAACAGCCAACATTAATTCCC	58.2			
	9.3	NGA63 F	AACCAAGGCACAGAAGCG	57.3	111	68	120
		NGA63 R	ACCCAAGTGATCGCCACC	59.6			
_	40.1	NGA248 F	TACCGAACCAAAACACAAAGG	56.1	143	129	no amplific.
		NGA248 R	TCTGTATCTCGGTGAATTCTCC	58.2			
	81.2	NGA128 F	GGTCTGTTGATGTCGTAAGTCG	1.09	180	82	no amplific.
		NGA128 R	ATCTTGAAACCTTTAGGGAGGG	58.2			
_	81.2	NGA280 F	CTGATCTCACGGACAATAGTGC	1.09	\$01	85	85
		NGA280 R	GGCTCCATAAAAGTGCACC	57.8			
_	111.4	NGAIII F	CTCCAGTTGGAAGCTAAAGGG	99	128	162	170
		NGAIII R	TGTTTTTAGGACAAATGGCG	70			
=	ca. 7.5	NGA 168 F	CCTICACATCCAAAACCCAC	57.8	213	217	208
		NGA 168 R	GCACATACCCACAACCAGAA	57.8			

•							
	07	19C11261	CGCTACGCTTTCGGTAAAG	57.8	161	661	961
	ç5	NGA 1126R	GCACAGTCCAAGTCACAACC	6.65			
		117047014	AAAGAGATGAGAATTTGGAC	51.7	114	120	114
=	62.2	NUASOIL	ACATATCAATATATTAAAGTAGC	49.5			
		NGA301K	ACAINICASININ				
			SUJULTA CITICA CITICA	59.6	151	135	135
	73	NGA168 F	וכחורוארוחרארוחרה				
		NGA168 R	GAGGACATGTATAGGAGCCTCG	6.19			
	77 63	AthBIO2 L	TGACCTCCTCTTCCATGGAG	59.9	141	209	139
=		A the los	TTAACAGAACCCAAAGCTTTC	54.5			
		Allibiot N					
		1 3110101111	AGGCAAATGTCCATTTCATTG	2.1	146	148	148
=	ca. 83	Athibiolic P	ACGACATGGCAGATTTCTCC	57.8			
		Alliobique n					
						761	140
111	3.4	NGA172 F	AGCTGCTTCCTTATAGCGTCC	33	162	05	0.51
		NGA172 R	CATCCGAATGCCATTGTTC	55.4			
	9 61	NGA126 F	GAAAAACGCTACTTTCGTGG	56.1	119	147	no amplific.
	0.77	NGA126 R	CAAGAGCAATATCAAGAGCAGC	58.2			
		2 671 4 614	CATGCAATTTGCATCTGAGG	55.8	107	68	no amplific.
=	17.5	NOA102 F		1.09			
		NGA162 K	CICIGICACICITICACICICA				

=	8.18	NGA6 F	TGGATTTCCTCCTCTTCAC	56.1	143	123	143
		NGA6 R	ATGGAGAGCTTACACTGATC	56.1			
٨	8.61	NGA12 F	AATGTTGTCCTCCCTCCTC	59.9	247	234	220
		NGA12 R	TGATGCTCTGAAACAAGAGC	58.2			
			-				
λI	24.1	NGA8 F	GAGGCCAAATCTTTATTTCGG	56.1	154	861	<u>8</u> 2
		NGA8 R	TGGCTTTCGTTTATAAACATCC	54.5			
) IV	102	NGA1107 L	GCGAAAAACAAAAAATCCA	50.2	150	140	140
		NGA1107 R	CGACGAATCGACAGAATTAGG	58			
>	8.11	NGA225 F	GAAATCCAAATCCCAGAGAGG	58	611	189	611
		NGA225 R	TCTCCCCACTAGTTTTGTGTCC	1.09			
							_
>	16.7	NGA249 F	TACCGTCAATTTCATCGCC	55.4	125	115	115
		NGA249 R	GGATCCCTAACTGTAAAATCCC	58.2			
>	6.61	CA72 F	AATCCCAGTAACCAACACACA	56.3	124	011	110
		CA72 R	CCCAGTCTAACCACGACCAC	6.19			
>	20	NGA151 F	GTTTTGGGAAGTTTTGCTGG	55.8	150	120	130
	·	NGA151 R	CAGTCTAAAAGCGAGAGTATGATG	58.6			
				-			

				1 (1)	651	123	130
>	24	NGA106 F	GTTATGGAGITTCTAGGGCACU	-00			
		NGA 106 R	TGCCCCATTTTGTTCTCTC	82.8			
		2000					
3	27.0	NGA139 F	AGAGCTACCAGATCCGATGG	59.9	174	132	761
>	07.0	NC 4130 D	GGTTTCGTTTCACTATCCAGG	8.88			
		N CCI VON					
		•					
	5	MCA76 E	GGAGAAATGTCACTCTCCACC	 	231	> 250	300
>	2	a alvon	OCCACATTANGE CATTANGE	57.8			
		NGA76 R	AGGCATGGGAGACATTACC				
		10103112	CTCCACCAATCATGCAAATG	55.8	148	156	146
>	1.19	AIHSOISI L		5 53			
		ATHSO191 R	TGATGTTGATGGAGATGGTCA	23.7			
				- 5	177	179	172
>	81.7	NGA129 F	TCAGGAGGAACTAAAGIGAGG	00.1	-		
		NGA129 R	CACACTGAAGATGGTCTTGAGG	90.1			

CLAIMS

- 1. An isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide functionally involved in the DNA mismatch repair system of a plant.
- 2. A DNA molecule according to claim 1 wherein said polypeptide is homologous to a mismatch repair polypeptide of a yeast or of a human.
 - 3. A DNA molecule according to claim 1 wherein said polypeptide is homologous to AtMSH3 (SEQ ID NO: 19) or to AtMSH6 (SEQ ID NO: 31).
- 4. An isolated and purified polypeptide functionally involved in the DNA no mismatch repair system of a plant.
 - 5. A polypeptide according to claim 4 which is homologous to a mismatch repair polypeptide of a yeast or of a human.
- 6. An isolated and purified polypeptide selected from the group consisting of a polypeptide encoded by the gene AtMSH3 (SEQ ID NO: 18), a polypeptide encoded by the gene AtMSH6 (SEQ ID NO:30), polypeptides homologous to a polypeptide encoded by the gene AtMSH3 (SEQ ID NO: 18) and polypeptides homologous to a polypeptide encoded by the gene AtMSH6 (SEQ ID NO:30).
- 7. An isolated and purified DNA molecule comprising a polynucleotide sequence selected from the group consisting of (i) a sequence encoding a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence; and (ii) a sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant.
- 8. A DNA molecule according to claim 7 comprising a polynucleotide sequence encoding a polynucleotide capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence.
- 9. A DNA molecule according to claim 8 wherein said polynucleotide is capable 30 of interfering with the expression of a plant polynucleotide sequence is a sense polynucleotide, an antisense polynucleotide or a ribozyme.
 - 10. A DNA molecule according to claim 7 comprising a polynucleotide sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant.

- 11. A DNA molecule according to claim 10 wherein said polypeptide is homologous to AtMSH3 (SEQ ID NO: 19) or to AtMSH6 (SEQ ID NO: 31).
- 12. A DNA molecule according to claim 10 further comprising a regulation element capable of causing overexpression of said polypeptide in a cell of said plant.
- 13. A chimeric gene comprising:
- a DNA sequence selected from the group consisting of (i) a sequence encoding a polynucleotide capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence, and (ii) a sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant; and

at least one regulation element capable of functioning in a plant cell.

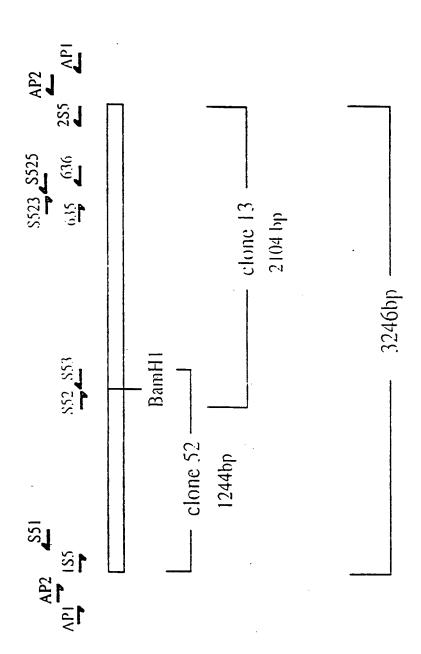
- 14. A chimeric gene according to claim 13 wherein said regulation element is selected from constitutive, inducible, tissue type specific and cell type specific promoters.
- 15. A chimeric gene according to claim 13 comprising a DNA sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant, wherein said regulation element is capable of causing overexpression of said polypeptide in a cell of said plant.
- 16. A chimeric gene according to claim 13 wherein said regulation element is selected from the group consisting of 35S, NOS, PR1a, AoPR1 and DMC1.
 - 17. A plasmid or vector comprising a chimeric gene according to any one of claims 13-16.
 - 18. A plant cell stably transformed, transfected or electroporated with a plasmid or vector according to claim 17.
- 25 19. A plant comprising a cell according to claim 18.
 - 20. A plant according to claim 19 selected from plants of the families Brassicaceae, Poaceae, Solanaceae, Asteraceae, Malvaceae, Fabaceae, Linaceae, Canabinaceae, Dauaceae and Cucurbitaceae.
- 21. A process for at least partially inactivating a DNA mismatch repair system of a plant cell, comprising transforming or transfecting said plant cell with a DNA molecule according to any one of claims 1-3 or 7-12 and causing said DNA sequence to express said polynucleotide or said polypeptide.
 - 22. A process for at least partially inactivating a DNA mismatch repair system of a plant cell, comprising transforming or transfecting said plant cell with a chimeric gene

according to any one of claims 13-16 and causing said DNA sequence to express said polynucleotide or said polypeptide.

- 23. A process for at least partially inactivating a DNA mismatch repair system of a plant cell, comprising transforming or transfecting said plant cell with a plasmid or vector according to claim 17 and causing said DNA sequence to express said polynucleotide or said polypeptide.
- 24. A process for increasing genetic variation in a plant comprising obtaining a hybrid plant from a first plant and a second plant, or cells thereof, said first and second plants being genetically different; altering the mismatch repair system in said hybrid plant; permitting said hybrid plant to self-fertilise and produce offspring plants; and screening said offspring plants for plants in which homeologous recombination has occurred.
- 25. A process according to claim 24 wherein a first gene is incapacitated in said first plant, a second gene is incapacitated in said second plant, and said first and second genes are incapacitated in said hybrid plant thereby altering the mismatch repair system of said hybrid plant.
 - 25. A process according to claim 25 wherein said incapacitation of the mismatch repair system of said hybrid plant is reversible.
- 26. A process according to claim 24 wherein a new genetic linkage of a desired characteristic trait or of a gene which contributes to a desired characteristic trait is 20 observable in at least one of said offspring plants.
- 27. A process for obtaining a plant having a desired characteristic, comprising altering the mismatch repair system in a plant, cell or plurality of cells of a plant which does not have said desired characteristic, permitting mutations to persist in said cells to produce mutated plant cells, deriving plants from said mutated plant cells, and screening said plants for a plant having said desired characteristic.
- 28. A process according to claim 27 wherein said step of altering the mismatch repair system comprises introducing into said hybrid plant, plant, cell or cells a chimeric gene according to claim 13 and permitting the chimeric gene to express a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence in a mismatch repair gene of the hybrid plant, plant, cell or cells, or a polypeptide capable of disrupting the DNA mismatch repair system of the hybrid plant, cell or cells.
 - 29. A process according to claim 28 comprising inactivating an MSH3 gene and/or an MSH6 gene of said plant.
- 30. A process according to claim 28 comprising inactivating an MSH3 gene and an MSH6 gene of said plant.

- 31. A process according to claim 27 comprising at least partially inactivating the mismatch repair system of said plant in a predetermined cell type or in a predetermined tissue of said plant.
- 32. A process according to claim 31 further comprising restoring mismatch repair 5 in said cell type or said tissue.
 - 33. An oligonucleotide capable of hybridising at 45°C under standard PCR conditions to a DNA molecule according to claim 1 with the proviso that said oligonucleotide is other than SEQ ID NO:1 or SEQ ID NO:2.
- 34. An oligonucleotide capable of hybridising at 45°C under standard PCR to conditions to the DNA of SEQ ID NO: 18 with the proviso that said oligonucleotide is other than SEQ ID NO:1 or SEQ ID NO:2.
 - 35. An oligonucleotide capable of hybridising at 45°C under standard PCR conditions to the DNA of SEQ ID NO:30 with the proviso that said oligonucleotide is other than SEQ ID NO:1 or SEQ ID NO:2.

Figure 1



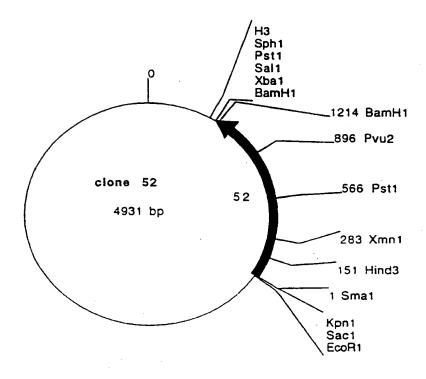


Figure 2

Comments/References: 52= 3' side of S5 (AtMSH3) 1244bp in pUC18/Sma1

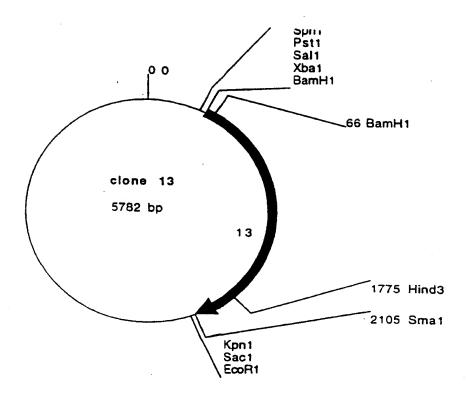


Figure 3

Comments/References: 13 = 3' side of S5 (AtMSH3) 2104bp in pUC18/Sma1

•	•	•	
	a)	
1	Ĭ	į	
	Ē	3	
		į	
ľ	-	ì	

80	144 15	204 35	264 55	324 75	384 95	444 115	504 135	564 155	624 175	684 195	744	804 235	864 255	924 275
AGCGCGCGAAAATTGGCAACCCCAAGTTCGCCATAGCCACGACCACGACCTTCCATTTCTCTTAAACGGAGGA	CCC P	CCG P	CTC L	GAT D	GTT V	A.A.G K	GGA G	TTC F	GCA A	GCA A	GAA E	AAT N	GAA E	GTT V
ACGG	GCT A	CCA P	CAC H	ည မ	TAT Y	CTA L	TTC	NAT N	AAT N	GGT G	CTT L	AGT S	ATT I	GTT V
TTAA	TTC	6	GAC D	GTA V	GAA E	GAG	TTC	CAC H	GTG V	CAT H	ACG T	CAG Q	GGT G	GAA
rcrc	TTC	ACA T	TCC	CCA P	GAG E	GTG V	AGA R	GAT D	CTG L	TCC	OCC A	TCA	TGT C	GGT G
CATT	CGT		CTC	AAC	200 P	GTG V	TAC Y	ATG M	AGA R	AAG K	A AAA	GGT	၁၅၅	ACA T
CTTC	TCT	TCA	CTT	CAA	TCG	CAA	AGG R	CAT H	AGA R	ATT I	ACC T	TTT F	TTA L	TCG S
CGAC	ATT	GAA E	AAG K	ACT	CCC P	CAG Q	TAC Y	GCT A	GTG V	gcc A	TAT Y	GGT G	ACA T	AT'T I
ACCA	ACG T	000 4	CGT R	CAC H	GAA	GW E	GGT G	TAC Y	CAT H	GCA	TTG L	GAA E	GAG E	GAA E
CACG	CAG	GT'A V	AAG K	CCT P	CTG L	TTG L	GTT V	AT'I I	TTC F	ACT T	gcg A	GAA E	TCG	GTT V
TAGC	CAG Q	CCG P	TCC S	TCT	TTT F	CCA P	GAA E	GGT	AAT N	GAA E	TCG S	GGT G	AAG K	ე ე
GCCA	A.A.G K	A.A.T. N	CCT	CTT L	AGA R	ACA T	GTG V	TTG L	TTG L	ACT T	CTG I.	GGT G	GTT V	GTT V
GTTC	CAA	CCG	TCT	AAA K	CAG 0	TAC Y	ATG M	GTG V	CGA R	CAG	GCA	TGT C	AGA R	GTT V
CCAA	A.A.G K	GAA	TTC F	CCT	CTC L	X X	TTG L	ეეე ა	TTT F	AAG K	CGG R	GGT G	GAG E	GGT G
CAAC	၁၅၅	CAC H	TCC S	AAG K	TTT E	AGG R	GTT V	GCA	ACA	GTG V	TTC F	GGT G	GAT D	GTC V
${ t r}$	ATG M	ACT T	GTA V	AAA K	AGA R	TCG S	GTG V	GCA A	CCA	GTA V	TTT F	AGT S	GTG V	AGA R
aaaa		CCG P	ACT	CCC	CAA	TCA S	GAT D	ATC I	GTG	GGT G	CCT P	ATA I	GTT V	GTT
ອວອວ	AGCA	TCC	9 8 8	TCA S	CAC H	TCA S	CCA P	GAG	AGT S	ATT I	၁၅၅	GAT D	TGT C	GAT D
AGCG	ATAAAGCAATT	X A		gcc A	TTA L	ACG T	TAC Y	GCG A	GCG A	A.A.G K	ACC	GAG E	GTT V	TTT F
	~	CCC P	ATA I	ეე 9	AAT N	GA.A	AAG K	GAC	ACG	TAC Y	იც გ	GCT A	TTG	AGT S
CCTAAGA	GATTACG	AAA K	AAG K	9 8 8	200 d	႕ ၁၁၁	AGC S	GAA E	ATG M	GGA G	AAC N	GCG A	T'IC F	ATG
-	81	145 16	205 36	265 56	325 76	385 96	445 116	505 136	565 156	625 176	665 196	745 216	805 23 6	865 256

925 276	TAT Y	GAA E	GAG E	TTC F	AAT	GAT D	AAT	TTC	ATG M	AGA R	AGT (GGA '	TTA (GAG	GCT (GTG ,	ATT	TTG	AGC S	TTG L	984 295
985 296	TCA S	CCA P	GCT	GAG	CTG L	$\frac{\Gamma\Gamma G}{L}$	CTT	299	CAG	CCT	CTT	TCA (CAA	CAA	ACT (GAG	AAG	TTT	TTG	GTG V	1044 315
.045 316	GCA A	CAT M	GCT A	GGA G	CCT P	ACC T	TCA	AAC	GTT V	CGA R	GTG (GAA (CGT (GCC	TCA (CTG (GAT D	TGT	TTC	AGC S	1104 335
105 336	AAT	GGT	AAT	GCA	GTA V	GAT	GAG	GTT ,	ATT	TCA	TTA '	TGT (GAA E	AAA K	ATC I	AGC S	GCA (GGT	AAC N	TTA L	1164 355
.165 356	GAA	GAT D	GAT D	X A	GAA E		AAG	CTG	GAG	GCT	GCT	GAA E	A.A.A.			TCT S	TGC	TTG L	ACA T	GTT V	1224 375
.225 376	CAT H	ACA	ATT I	ATG M	AAC N	ATG	CCA P	CAT	CTG L	ACT	GTT	o CAR	225		200	CTA	ACG	TTT F	7GC C	CAT H	1284 395
.285 396	CTC L	A A	CAG Q	TTT F	GGA G	TTT F	GAA	AGG R	ATC 1	CTT	TAC	CAA	999	229		TTT (CGC	TCT	TTG L	TCA S	1344 415
345 416	AGT S	AAC	ACA T	GAG E	ATG M	ACT	CTC L	TCA	200	AAT	ACT	CTG (CAN	CAG	TTG (GAG (GTT	GTG V	A A	AAT N	1404 435
405	AAT N	TCA S	GAT D	GGA G	TCG S	GAA E	TCT	200	TCC	TTA L	TTC	CAT H	AAT	ATG M	AAT	CAC H	ACA T	CTT	ACA T	GTA V	1464 455
1465 456	TAT Y	GCT	TCC S	AGG R	CTT L	CTT L	AGA R	CAC	TGG	GrG V	AC'F T	CAT (CCT	CTA	TGC (GAT D	AGA R	AAT N	TTG L	ATA I	1524 475
525 476	TCT	GCT A	CGG R	CTT L	GAT D	GCT A	GTT V	TCT	GAG	ATT I	TCT	GCT A	TGC	ATG	66.A 5	S	CAT	AGT	TCT	TCC	1584 495
585 496	CAG Q	CTC L	AGC S	AGT S	GAG E	TTG L	GTT V	GAA	GAA	GGT	TCT	GAG	AGA (ATT	GTA V	TCA	CCT	GAG E	TTT F	1644 515
645 516	TAT Y	CTC L	GTG V	CTC L	TCC S	TCA S	GTC V	${f r}{f r}{f G}$	ACA T	GCT	ATG	TCT	AGA .	TCA	TCT (GAT	ATT (CAA	CGT R	GGA	1704 535
.705 536	ATA I	ACA	AGA R	ATC I	TTT F	CAT H	000 8	ACT	GCT	A A	200	ACA T	GAG '	TTC	ATT	GCA (GTT V	ATG M	GAA	GCT A	1764 555
765 556	ATT I	TTA L		928 4	999 9	AAG K	CAA	ATT I	CAG Q	cgg crr R L igure		ပ်) ၁၉၆၆	C ATA AAG I K (Continu	. 0	CAA O	GAC	S	GAA	ATG M	AGG R	1824 575

1884 595	1944 615	2004 635	2064 655	2124 675	2184 695	2244 715	2304 735	2364 755	2424	2484 795	2544 815	2604 835	2664 855	2724 875
es.	GCT A	GCT A	TTT F	TTG L	AAG K	GCA A	AGT S	CAC H	TGT C	CAA	ATC I	ATA I	GGT G	GAA E
ATT TO	GCG G	CTT G	TCA T S F	CAT T H	ACC A	CTA G	TTC P	TTG C	GAC 1	TTA (ATT I	TCC	GAT	CTA
GTT /	GAA	GAG .	GCT	ACA	AGC S	GCT A	AGT S	TGT C	GAT D	ATA I	CAA	ATT I	CTT	F
TCT	A.A.G K	CCT P	ATA I	ATC I	AAT N	CTA L	AAG K	GAC D	GTG V	ACT	16C	TTA L	GTG V	T T
ATT I	AAT	TTT F	TCG S	9 999	GTA V	GAG E	CTC L	CTG L	TTT	GAG	, TAT Y	GCT A	CAC H	A AGT S
TTG L	CTA L	CAA O	TCC S	TCG S	KAA K	GAT D	TTC F	GCA A	GAG	CTG	GAA E	A GTT V	3 CTG	C AGA R 1 d)
A AA	9 8 8	GAC	GAT U	G1'G V	GTG V	T FG L	AGT S	GCT A	CCC P	GTA	999 1	CAA	AAG K	NG CAT GGC H G
AGA R	TCT	AGC S	CTG L	CAA	TGG W) (0.00) (0.00 (0.00) (0	GAT D	CTT L	CGT R	CCT P	GAA	CGT R	3 € 5 6 7	Cont
TTG L	CTC L	TCC	AAG K	CT'I L	AAT N	GCT A	TGG ™	GCT A	GTC V	CAT H	GCA A	ATC	TTC	ပီဝ
CTT L	CTT L	AC'T T	GAA E	TTT F	ATG H	GTA V	TCG	CAA	TAT Y	CGT R	CAT H	TAT Y	TCA S	ATC I
ACT	AAA K	A'I'C I	AGG R	GAA E	CCT P	A'fa I	GCT A	GTT V	AAC	GGT	TTG L	7GC C	929 •	AGT A'S IS
TCT	GGA	CTA L	ATC I	TTG	GTC V	GAA E	CGA R	GCC	AAG K	TCT	ATT I	AGC S	CCA P	GAC D
CGA R	9 8 8	ATA I	GTC	AAT	AAG K	CCA P	AAC	GCT A	AAC	c.AG	ACA T	A.A.G. K	GTA V	TCA
GTG V	AAT N	GAC D	TTA I.	CGA R	TCC	၁၁၁ ၂	GTG V	A.A.G.	AGA R	ATA I	GAC D	GGA G	TTT F	GCT
ACT	GAC	CTC L	GTT V	ATT	GAT D	CAT H	ATT I	TTT F	TCT		AAT N	GGA G	TCC S	GGT
GCA A	GTT	TTG L	GCA	GCT.	GTT V	TAT Y	200 V	GAT D	CTA		CCA P	ATG M	GGT G	ATG M
TCT	GTG V	GAC D	CAA	CTC L) 	CGA R	CTT		ACT T	GAG	GTC V	AAC N	GTT V	CGG R
C A		GGT	CGC R	A AG	CTG L	ATT I	CAT H	TAC Y	TCA	GTT V	TTC F	CCT	CAG O	
ATG				A.A.G	GAG E	ACT T	GAA	TAC Y	CTT L	CCA	AAC	GGA	GCT A	
AGT 1					ATA I	AAG K			TCC	GAA	GAT	ACC	ATG	GTT
1825 576		1945 616	2005 636	2065 656	2125 676	2185 696	2245	2305	2365	2425	2485	2545	2605	2665

2725 876	GAA E	TTA L	AGT S	GAA	GCG A	TCA (CAC	ATA	ATC	AGA R	ACC	TGT C	rct 1	rcr (CGT R	TCG (S	CTT	GTT 1	ATA	TTA L	2784 895
2785 896	GAT D	GAG E	CTT L	GGA G	AGA R	ວອອ	ACT ,	AGC	ACA	CAC	GAC	GGT (GTA (gcc A	ATT	000 4	TAT	GCA	ACA T	TTA L	2844 915
2845 916	CAG O	CAT H	CTC	CTA L	GCA	GAA	AAG K	AGA R	TGT	TTG L	GTT V	CTT	TTT (GTC V	ACG	CAT '	TAC	CCT	GAA	ATA I	2904 935
2905 936	GCT	GAG E	ATC I	AGT S	AAC	GGA	TTC	CCA P	GGT	TCT	GTT	, 5 5 5 5 5 5	ACA T	TAC	CAT H	GTC '	TCG	TAT	CTG	ACA T	2964 955
2965 956	TTG	CAG	A.A.G		A AA	၁၅ ၅	AGT	TAT Y	GAT D	CAT H	GAT	GAT	GTG ,	ACC	TAC	CTA	TAT	AAG K	CTT	GTG V	3024 975
3025 976	CGT R	GGT G	CTT L	76C C	AGC S	AGG	AGC S	TTT	GGT G	TTT F	AAG K	GTT V	GCT A	CAG	CTT L	000 8	CAG	ATA I	CCT	CCA	3084 995
3085 996	TCA	TGT C	ATA I	CGT R	CGA R	225 4	ATT I	TCA	ATG M	GCT A	GCA A	AAA K	TTG	GAA	GCT	GAG	GTA V	CGT	GCA	AGA R	3144
3145 1016	GAG	AGA	AAT	ACA T	CGC R	ATG M	GGA G	GAA E	CCA P	GAA	GGA	CAT H	GAN E	GAA	၁၁၁ ၂	AGA R	၁ဗ္ဗ	GCA	GAA	GAA E	3204 1035
3205 1036	TCT	ATT I	TCG S	GCT A	C'I'A L	GGT G	GAC	${ m TTG}_{ m L}$	TTT F	GCA A	GAC D	CTG L	§×	T'IT F	GCT A	CTC	TCT	GAA	GAG	GAC D	3264 1055
3265 1056	CCT P	TGG W	X X	GCA	T'IC F	GAG	TTT	TTA L	AAG K	CAT' H	GCT A	TGG ₩	AAG K	ATT 1	GCT	၁၁၁	AAA K	ATC	AGA R	CTA L	3324 1075
3325 1076	A A	CCA P	ACT T	TGT C	TCA S	TTT F	TGA	TTT	WTCI	rtaac	CATTA	TGA TTTAATCTTAACATTATAGCAACTGCAAGGTCTTGATCATCTGTTAGTTGCG	AACT	GCAA	GGTC	TTGA	TCAT	стст	TAGT	rccc	3397 1082
3398 1	TAC	TACTAACTT		ATG TC	TGT AT	ATT AG I S	AGT AT S I	ATA AC	ACA AGA AAA C T R K E	SA AZ K	\$ ₹	GAG AAT TAG AGAG ATG	T. TA	DV 91	AG A		GAT T D S	TCT A	AAT C	6	3458 5
3459 6	GTG V	TTG L	0.0	AG TAC	ATC	TTT	TCT	CCA)) 4	GCA A i gu	TAA re 4	C GCA TAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAA onti	AAAA inue	AAAA	AAAA	AAAA	AAAA	¥		3522 16

e amai,

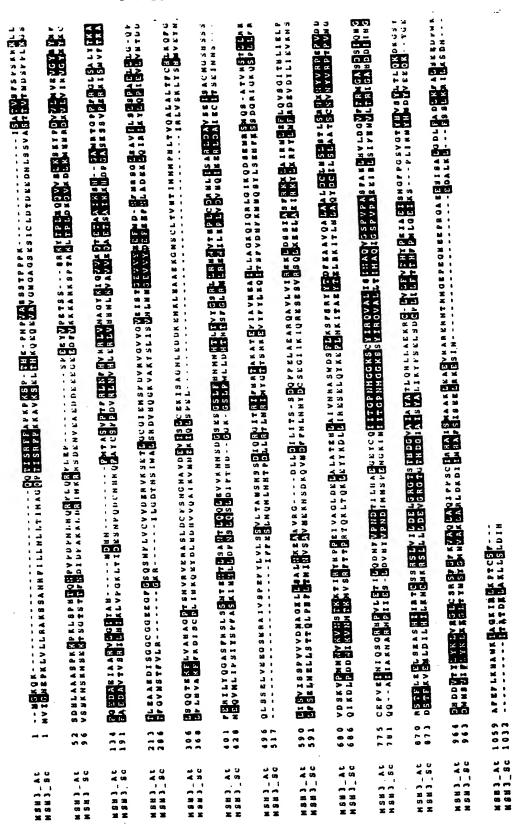
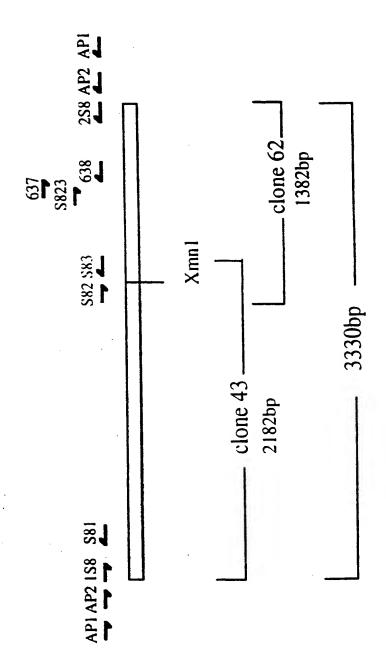


Figure 6



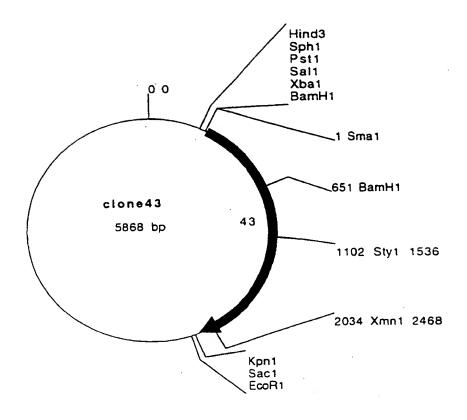


Figure 7

Comments/References: 43= 5' side of S8 (AtMSH6) 2182 bp in pUC18/Sma1

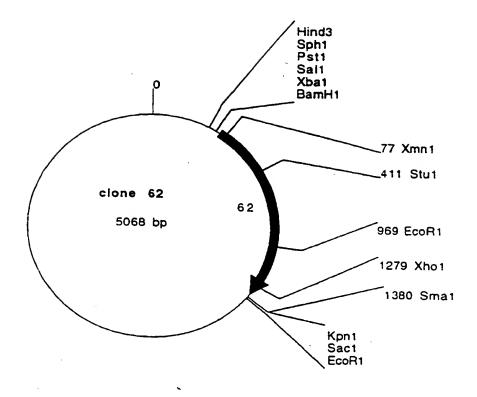


Figure 8

Comments/References: 62= 3' side of S8 (AtMSH6) 1379bp in pUC18/Sma1

0
н
3
ъ
٠Ă
بعا

80	153	213 24	273	333 64	393 84	453 104	513 124	573 144	633 164	693 184	753 204	813 224	873 244	933 264
သသ	_O	TCC	999 999	AGA R	CCG P	AAA. K	GAT D	X AA	GAA	GCT A	GTA V	GAA	AGA R	TTC
rTCC	CAG	GTT	GAA E	GTT V	AAG K	GTA V	AAT N	GGT G	GTA V	CGT R	CCT P	¥ A	AAT N	GTT V
TCT	s cGC R	TTG (AAG (R	GAG E	TTT F	TTT F	CTG L	AAT N	TCA S	CCA P	GTT V	AAG K	GCC	GAT D
CCA	c cAG	GGT 1 G	GTG /	GAT O	GGA G	A.A.G K	CCG P	AAT N	AGA R	CGT R	AAG K	GAG E	GAT D	CCT P
0000	A ATC	AAG O	AAT O	GTC (S	CAT	GTT V	TCC S	CTT L	ATG M	GAT D	GGA G	AGG R	CCA P
TTT	CTCTCAGCTCAAAACATCGTTTCTCTCTCTCTCTCTCACAATTCCAAAAA ATG	ACG 7	TTT /	TCT (S	000 d	ATG M	GTT V	CGT R	GAA	9 999	GAG E	TGT	ATC I	ATA I
AAA1	TCC	ACT A	CGA 1	AAA K	CTG (ATT	GAT D	۲۰۱۰٬۲ آ	GCT	CCA	A.A.G K	GTT V	CGA R	CAC
GCGA	CAA1	6CG <i>H</i>	CCA (TCG 1	GTC (AAT	GAA	CAA	AGA R	ACA T	TTT F	000 P	TCT	TTA L
CAAG	CTCA	CCG O	GGA C	GTT 7	CGT (TCC S	CGA R	CCT P	ე ე	GAA E	ACT T	GAT D	TCT	ACC
GACG	CTCI	ACC (T	0 400 0	GCT (CGT	TTC	AGC	AT'T I	AGT S	CCA P	ATG	CAG Q	GAG E	AAG K
CTAC	CACT	CCC F	AGC (TTT (CTG L	AGG R	GTrf V	TTC F	GGT G	GAA E	CTC L	CTT L	AGA R
ATTT	יכדכו	AAA O	; 9	CGT '	GTT (TCC S	GAG	GA'T D	AGT S	CCT P	GAT D	ATG	TGG ¥	GAT' D
0000	TTC	CAA A) 0 0 0 0 0 0	GTA (AAG K	TCG	GGA G	Z AST	TTT F	GTT V	gag E	A AA	GAA E	TAC Y
TTT	TCGI	TTC () 999	TCT (GAG I	GCT	TCT	GCT A	GCT A	GAT D	CTG L	CTG L	TTT F	CTT L
ATCG	AACA	TTC 1) 299	GCT 3	CCG (GAT	TGT	AAG K	CAT H	ეეე ეეეე	GTT V	AGG R	A AA	222 6
GAGT	TCA	TCT 1	AGC (S	GAC (CCA (GST	GAT		A.A.C.	GAT D	CGA R	A A	ACC T	GAT D
TGAG	CAGO	TTG 1	GCT P) 299	ACT (225	CGA R	TGT	AGA R	GTA V	AAG K	AAC N	GGA G	GAT D
သသ	CTCT	ATT T I	GCT G	AAA K		TCC (GAT (CTA	GAA E	GGA	TTG L	TCT	GAA	CCT
TTGA	AATT	TCG A		GCT P		GAA						GAC	AAC	CGT R
aaaagttgagccctgaggagtatcgtttccgccatttctacgacgcaaaagcgaaaattttggcgccaatctttcccccc	TTTCGAAT1	AGA T R S				GCT (TCA	ACT (GAT				AGA R
1. A	-													

934 265	AAG 7	AAA ATG K M	-	င်္သ	GCA T	rca c s	CPA A	AAG C	CAA T	TAT T Y	rgg A	AGT G	GT'F A	AAG A(AGT G	GAA TI	TAT A'	ATG G M	GAC A	ATT I	993 284
9 5	GTG V	ប្រធ	T TTC	TTT F	AAA K	GTG V	999	AAA K	TTT F	TAT	GAG	CTG	TAT (GAG	CTA	GAT (GCG	GAA	TTA L	GGT G	1053 304
4 N	CAC H	₹ϫ		CTT L	GAC	TGG *	AAG K	ATG	ACC	ATG	AGT S	GGT	GTG (GGA G	AAA	TGC /	AGA (CAG	GTT V	GGT	1113 324
4 N	ATC I	ည္	GAA E	AGT S	999 999	ATA I	GAT D	GAG	GCA A	GTG V	CAA	AAG	CTA	TTA (GCT A	CGT (GGA	TAT	A A A	GTT V	1173
4 N	GGA	CGA R	ATC I	GAG E	CAG 0	CTA	GAA E	ACA	TCT	GAC				GCC A	AGA (GGT (GCT	AAT		ATA I	1233 364
4 2	ATT I	CCA P	AGG R	AAG K	CTA L	GTT V	CAG Q	GTA V	TTA	ACT	CCA	TCA	ACA (T	GCA	AGC S	GAG (GGA G	AAC	ATC I	9 999	1293 384
ه	CCT	GAT D	9 9	GTC V	CAT H	CTT	CTT L	GCT.	ATA	AAA K	GAG	ATC	A X	ATG (GAG (CTA (CAA			TCA	1353 404
A 70	ACT	GTG V	TAT Y	GGA	TTT F	GCT A	TTT F	GTT	GAC	TGT	GCT	9 8	TTG /	AGG 7	TTT '	TGG (GTT (999	TCC	ATC . I	1413 424
4 N	AGC S	GAT D	GAT D	GCA	TCA S	TGT	GCT A	GCT.	CIT.	GGA G	GCG A	TTA '	TTG /	ATG (CAG (GTT 7	TCT	CCA	AAG K	GAA E	1473 444
4.N	GTG V	TTA 1	TAT Y	GAC D	AGT S	X AAA		CTA L	TCA S	AGA R	GAA	GCA (CAA 2	AAG (GCT (CTA A	AGG R	AAA K	TAT Y	ACG T	1533 464
4. N	TTG L	ACA T	9 9	TCT	ACG	900 4	GTA	CAG O	TTG (GCT A	CCA (GTA (CCA (CAN C	GTA I	ATG (999	GAT D	ACA T	GAT D	1593 484
44 YO	GCT	GCT A		GTT	AGA R	AAT	ATA	ATA (GAA :	TCT	AAC (GGA	TAC 1	TTT /	A A A	GGT 1	TCT S	TCT (S	GAA E	TCA S	1653 504
4 N	TGG W	AAC N		GCT A	GTT	GAT	GGT	CTA L	AAT (GAA '	TGT (GAT (GTT C	GCC (CTT /	AGT (GCT (CTT (GGA	GAG E	1713 524
44 1V	CTA L	ATT I			CTG	TCT	AGG R	CTA /	AAG (CTA (GAA C	GAT (GTA O	CTT A	AAG C	CAT G	999	GAT /	ATT	TTT F	1773 544
1774 545	CC A	TAC Y		GTT V	TAC	AGG R	GGT	TGT	CTC 1	AGA ATT R I Figure		GAT (D) (C) (C)	GGC (G)	CAG # 0 1 nued)	ACG A	ATG G M	GTA A	AAT O	CTT	GAG	1833 564

1893 584	1953 604	2013 624	2073 644	2133 664	2193 684	2253 704	2313 724	2373 744	2433 764	2493 784	2553 804	2613 824	2673 844	2733 864
AAC N	GAT D	AGT S	CGC R	GTG V	GAT D	AAA K) GCC	ACT T	. CAC H	ogy G	ACA T	r GGT G		r cgr R
GAT D	A A	gaa E	GGA G	X AAA	ATT I	7GT C	GCA A	GAA	ATT I	GCT A	AAA K	GAT D	AGC	CTT
CTT	CTC L	TCA S	CTC L	X X	GGA G	CTC L	GAA E	GCT A	GTC V	TCT	CAG O	9 9	၁၅၅	CTT L
TAT Y	CCA P	AAC N	CTG L	9 9 9	AGT S	A A	TTC F	AAC	GAG E	CTC L	AAT N	GCA	AGT S	ACT
AAA K	CAT H	GCA A	AGA R	CTG L	AGA R	TAT Y	CAA Q	GAA E	TCT S	AGT S	CAG Q	GTT V	AGC	TCA
TAC Y	TGC	ACG T	GAA	Crr	TTC F	CTT L	TCT S	GAT D	1GG ₩	GCA A	GAT D	GCA A	AGA R	AAA K
TTG L	ATC I	TTC	T'ſÅ L	GCT A	9 6 6	TTG L	CTT L	ACA T	\$ 0 0	GCA	ACA T	TTT F	AGA R	GGA AA G K Inued)
ACC T	TGG W	GAA E	GAC D	CC.T P	K XX	N.T. S	TTT F	GTG V	ACT T	ATC I	GCT A	CCA P	GCT A	GGC G
999	AAT' N	gaa E	CCA P	TTG L	GTG V	VIQ W	TTA L	GAT D	GCA A	GCA A	GAA E	CAT H	GAG E	ATG M
TCA	AGG R	GTT V	CTT L	GTG V	ATT I	ATG M	GAG E	CAA O	AGA R	TTT F	TCA S	TGG ₩	၁၅၅	CCA AAC P N Figure
CCT P	T'fA L	GTA V	A X	TCT S	CAN	AA'I' N	CTA L	AAC	GAA E	TCT	GAA E	CTA L	CTT L	
GGT	CTC L	GAT D	CAC) (ე ეეე	TCA S	ე ეეე	CAG Q	ATC I	AGA R	ည ၂	GGA G	CTC	GGA
GGT	CGA	CTT L	CTC 1.	TCA	TTT F	GAA E	AGC S	TAT Y	TTT F	CTG L	TTT F	O CAA	ATA I	ACG T
GAT D	AAG K	CGG R	TAT Y	TCA	GC:A A	A AG	AAA K	AAT N	CTT L	GTC V	ATT	ATC I	GAT D	CTG
TGT	GGT	A AA	CAG	CGA R	A AA	CAG 0	GGA G	CCA P	gaa E	GAT D	GTT V	AAA K	AAT	TTA L
AGC S	ACT T.	AAT N	ე ეენ	GTT V	GTT V	CTA L	GTA V	TTT F	ATC I	CTA L	CCT P	CTT	0 0 0	TTG L
AAT	CCA	ATC I	ACT T	AGC S	CGA R	GCT A	TTA L	GAC	CTT L	76C C	AGG R	ATA I	GTT V	TCA
AAC	AGT	AGC		TCT	CAA	TTG L	ATA I	AGC S	ATA I	AGC S) (CC	CCA P		CGG R
TTT /		GAA			A AA	TTG L	CCT P	GAT	ACA T	ATA I	ATG M	999 9	TTG L	CCT P
ATA 1		GTA (CTG L	CTG L		ATA I	CTC	ACC T	AGC	AAA	o car	CAT H
.834 565		1954 605	2014 625	2074 645	2134 665	2194 685	2254 705	2314	2374	2434	2494	2554 805	2614 825	2674 845

2734 865	GCA A	ACA	TGT C	cTG L	225 ₹	GTT V	ATC	TTT	0 0 0 0 0	CAA	CTT	66C 1	TGC 1	TAC (GTG (, 500 P	rgr	GAG	TCT	7.00 C	2793 884
2794 885	GAA E	ATC I	TCC	CTC L	GTG V	GAT	ACT	ATC I	TTC	ACA I	AGG (CTT () 299	GCA ,	rcr (S	GAT	AGA R	ATC	ATG	ACA T	2853 904
2854 905	GGA G	GAG E	AGT S	ACC T	TŢT F	TTG L	GTA V	GAA	TGC	ACT (GAG E	ACA (T	; 925	TCA (GTT V	CIT	CAG Q	AAT	GCA A	ACT	2913 924
2914 925	CAG O	GAT D	TCA S	CTA L	GTA V	ATC I	CTT L	GAC	GAA	CTG (၁၁၁၁	AGA (GGA 2	ACT	AGT	ACT	TTC F	GAT D	GGA G	TAC Y	2973 944
2974 945	9 8 8	ATT I	GCA A	TAC Y	TCG S	GTT	TTT E	CGT R	CAC	CTG	GTA V	GAG 1	AAA Y	GTT	CAA	TGT	000 R	ATG M	CTC	TTT F	3033 964
3034 965	GCA	ACA T	CAT H	TAC Y	CAC H	CCT P	CTC L	ACC T	AAG K	GAN	TTC	000 4	TCT	CAC	CCA P	CGT R	GTC	ACC	TCG S	AAA K	3093 984
3094 985	CAC H	ATG M	GCT A	TGC C	GCA	TTČ F	AAA K	TCA S	AGA R	TCT	GAT D	TAT	CAA	CCA	CGT	GGT	TGT C	GAT D	CAA	GAC D	3153 1004
3154 1005	CTA L	GTG V	TTC F	TTG L	TAC Y	CGT R	TTA L	ACC T	GAG E	GGA G	GCT A	TGT	CCT P	GAG	AGC S	TAC Y	GGA G	CTT L	CAA	GTG V	3213 1024
3214 1025	GCA	CTC		GCT A	GGA G	ATA I	CCA P	AAC N	CAA	GTG V	GT'F V	GAA	ACA	GCA A	TCA S	GGT G	GCT A	GCT A	CAA	000 4	3273 1044
3274 1045	ATG	X A	S AGA R	TCA	ATT I	ეეე ე	GC;A	AAC	TTC F	AAG K	TCA	AGT. S	GAG	CTA L	AGA R	TCT S	GAG E	TTC F	TCA S	AGT	3333 1064
3334 1065	CTG L	EAT	GAA	GAC	TGG X	CTC L	A.A.G K	TCA	TTG L	GTG V	GGT	ATT I	TCT S	CGA R	GTC V	gcc A	CAC H	AAC	AAT	000 A	3393 1084
3394 1085	6	ATT	၁၅၅	GAA E	GAT D	GAC	TAC	GAC D	ACT	TTG L	T'I'T F	7GC C	TTA	TGG W	CAT H	GAG E	ATC I	AAA K	TCC S	TCT	3453 1104
3454 1105	TAC	TGT	GTT	CCC P	AAA K	TAA	ATG M	GCT	ATG M	ACA T	TAA	CACT	ATCT	GAAG	CTCG	TTAA	GTCI	TTT	CACTATCTGAAGCTCGTTAAGTCTTTTGCCTCTCT	TCT	3521 5
3522	Φ Σ	ATG T M	TTT A F I	ATT CO	CCT CI	CTT A	AAA AA K	AAA TC K C	TGC TT	TTA TA	TAT AT Y I	ATC AA I K	AAA AA K	AAA TTG K L	G TTT F	T CCT		CGA TT R L	TTA AP	AAA K	3579 19
3580 20	AAA K	K K	K K	K K	X AA	X X	X AAA	AAA K	AAA K	म् हां	Figure 9		(C ntinued)	penu	_						3606 28

=	
_	
<u>د</u>	
Ξ	
ដ	0
=	
_	

1	69 Perkyanylpropries and Dasglefinialisrvkvdördcsorfinas 7Veluboslemandy ipopanadytosenades og anales 108 popuminanosenistros Dasglefinialisrvkvdördcsorfinas og bellemandy ipopanadytosenades og de de de de	16) VEDIGVDODV POPETVORNÝM ABRLEÑV LEDEM FYRNÝM VOJENÍM KULÚĽ V COBÍM A VEROTES HUBBER BURÍM BODÍM TERMÍM PPOV PAĽA 210. lasttere kybym skrem samunámoskaka v voja kybym skrem skrem sím komeratom vos - 1000 mily modem kybym kym	260 arsq <u>kovisty kstymoty i kvydeb</u> de Dechareden en spyceruy gisty of dramater or byky and interparanted for the dechared of the second of the	17.2 VORSAGERATOR VON MERNINGER VON TVERS BELLE STATE OF THE SECOND STATE OF THE SECON	463 KYTLTOSTAVQLAPVEQVMODYDAROVAMINEDAONIKOSKISZNICAVDOLNECDZALNYLDREKAMEDARONIADVLEMODIFFROVINGCLAIOONIP 522 pmaarmaiynevaussytDübatyakiisiseliis yeilulekulastotokazooniyaineliaminalmanastorkaunamvidiillo	ses musicamos somo Mudacos somo destante de martino de mando de sus estantes de martinos d	918 SYNSS	125 IDBOFFM	803 SATROFIENTOORMENDERVELCOOLFVONDILLEGARANGE HORDE ENTRE NATURAL ENTRE NO DEVICESAVETE SALTE ENTRE NATURAL DE ENTRE NATU	POS SESTILIZACIONE AND UNIONE SENTIDA MANDEN POSTANTA SENTIDA NA CARROLA SENTENCIA DE LA SENTENCIA DEL SENTENCIA	1010 BLIEGO.C.BET D. ORGERT DE BOVVETTS ONNO
					٠	J			٠	د	
18-988814	15-9202	18 - 92 929 202 202 202 202 202 202 202 202 2	- 9 X 9 X 9 X 9 X 9 X 9 X 9 X 9 X 9 X 9	16 - 9.2 9.2 H	A - B H A A A A A A A A A A A A A A A A A A	14 - 92 0 X 3 Y	14 - 92 92 92 92 92 92 92 92 92 92 92 92 92			ALEGMA-91	200

	50
TTTTTTGGTTGCTAACAATAAAGGTATACGGTTTTATGTCATCAATATAA	100
CTATATATAAAAGAAATGAAAGATATATTATTTTTTTTTT	150
AAAACAACAAGACTTTTTTTTTTTTTTTTTTTTTTTTTT	200
GATAAACGACATCGTTTAATCATTTCCCAATTTTACCCCTAAGTTTAACA	250
CCTAGAACCTTCTCCATCTTCGCAAGCACAGCCTGATTAGGAACAGCTTT	300
ACCATTCTCATATTCCTGAACTACCTGAGTCCTCTCATTGATCTGTTTCG	350
CCAAATCCGCTTGTGACATCTTCTTCTCCAATCTCGCTTTCTGTATCATC	400
AACCTCACCTCTGCTTTCACACGATCCATCGCCGCAGGCTCTGTTTCTTC	450
TTCCAGCTTCTTCGTGTTAATCACCGGAACCGCCGTAGATTTCCCCTTTT	500
TGTTCGAACCGGCATCGAATTTCTTAACCGTTTGAACCGCGACACCGTTT	550
CTCAGAGCTGCGTTAACCGCTTTCGGATCGCGTAGGTCTTGGCTCTTTTG	600
TTTTGATTTGTGGAGAACTACTGGTTCCCAGTCTTGTGTTACTGCTCCTG	
GGTATCTGCTCGGCATCGATGAATTGAGAGAAAGGAACAACGCGAAA	650
ATTTTATTAATCTGAGTTTTGAAATTGAGAAACGATGAAGATGAAGAATG	700
TTGTTGAGAGGATTGTGATATTATATATACGAAGATTGGTTTCTGGAGA	750
ATTCGATCATCTTTTCTCCATTTTCGTCTCTGGAACGTTCTTAGAGATG	800
ATTGACGACGTGTCATTATCTGATTTGCAGTTAACCAATGCTTTTTGGGT	850
TGGATTCGTGGTACACCATATTATCCGATTTGGCTCAATGGTTTTATATA	900
AATTTGGTTTTCGGTTCGGTTATGAGTTATCATTAAAATTAAGCTAACCA	950
AAAATTTTCGTAAAATTTATTTCGGTTTCAATTCGGATCCCTTACTTCCA	1000
GAACCGAATTATTCGAAACCGGGGTTAGCCGAACCGAATACCAATGCCTG	1050
ATTGACTCGTTGGCTAGAAAGATCCAACGGTATACAATAATAGAACATAA	1100
ATCGGACGGTCATCAAAGCCTCAAAGAGTGAACAGTCAACAAAAAAAGTT	1150 1200
GAGCCCTGAGGAGTATCGTTTCCGCCATTTCTACGACGCAAAGGCGAAAAT	1250
TTTTGGCGCCAATCTTTCCCCCCCTTTCGAATTCTCTCAGCTCAAAACATC	1300
GTTTCTCTCTCTCTCTCTCACAATTCCAAAAAATGCAGCGCCAGAGAT	1350
CGATTTTGTCTTCCTAAAAACCCACGGCGGCGACTACGAAGGGTTTG	1400
GTTTCCGGCGATGCTGCTAGCGGCGGGGGGGGGCAGCGGAGACCACGATTT	1450
AATGTGAAGGAAGGGGATGCTAAAGGCGACGCTTCTGTACGTTTTGCTGT	1500
TTCGAAATCTGTCGATGAGGTTAGAGGAACGGATACTCCACCGGAGAAGG	1550
TTCCGCGTCGTGTCCTGCCGTCTGGATTTAAGCCGGCTGAATCCGCCGGT	- :
GATGCTTCGTCCCTGTTCTCCAATATTATGCATAAGTTTGTAAAAGTCGA	1600
TGATCGAGATTGTTCTGGAGAGAGGTACTAATCTTCGATTCTCTTAATTT	1650
TGTTATCTTTAGCTGGAAGAAGATTCGTGTAATTTGTTGTATTCGTT	1700
GGAGAGATTCTGATTACTGCATTGGATCGTTGTTTACAAATTTTCAGGAG	1750
CCGAGAAGATGTTGTTCCGCTGAATGATTCATCTCTATGTATG	1800
ATGATGTTATTCCTCAATTTCGTTCCAATAATGGTAAAACTCAAGAAAGA	1850
AACCATGCTTTTAGTTTCAGTGGGAGAGCTGAACTTAGATCAGTAGAAGA	1900
TATAGGAGTAGATGGCGATGTTCCTGGTCCAGAAACACCAGGGATGCGTC	1950
CACGTGCTTCTCGCTTGAAGCGAGTTCTGGAGGATGAAATGACTTTTAAG	2000
GAGGATAAGGTTCCTGTATTGGACTCTAACAAAAGGCTGAAAATGCTCCA	2050
GGATCCGGTTTGTGGAGAGAAGAAGAAGTAAACGAAGGAACCAAATTTG	2100
AATGGETTGAGTCTTCTCGAATCAGGGATGCCAATAGAAGACGTCCTGAT	2150
GATCCCCTTTACGATAGAAAGACCTTACACATACCACCTGATGTTTTCAA	2200

GAAAATGTCTGCATCACAAAAGCAATATTGGAGTGTTAAGAGTGAATATA	2250
GAAAATGTCTGCATCACAAAAGCAATATTGGAGTGTTACTAATCTAGTGTT	2300
TGGACATTGTGCTTTCTTTAAAGTGGTTAGTAACTATTAATCTAGTGTT	2350
CAATCCATTTCCTCAATGTGATTTGTTCACTTACATCTGTTTACGTTATG	2400
CTCTTCTCAGGGGAAATTTTATGAGCTGTATGAGCTAGATGCGGAAAATGC	2450
GTCACAAGGAGCTTGACTGGAAGATGACCATGAGTGGTGTGGGAAAATGC	2500
AGACAGGTAAATTAGTTGAAACAACTGGCCTGCTTGAATTATTGTGTCTA	2550
TAAATTTTGACACCACCTTTTGTTTCAGGTTGGTATCTCTGAAAGTGGGA	2600
TAGATGAGGCAGTGCAAAAGCTATTAGCTCGTGGGTAAGGGAACCATCAT	2650
ACTTTATGGAATTCGTTTACTGCTACTTCGGCTAGGATTTAAGAAATGGA	2700
AATCACTTCAAGCATCATTAGTTAGGATCCTGAGAACTCAGGATGTTTTC	2750
TTATTCGTTATATAATAAGTCTTTTCATCAAGGAGTAACAAAACTT	2800
GCACAATATTTGTGTGCTCACTGGCAAGGCATATATACCCAGCTAACCTT	2850
TGCTAGTTCACTGTAGTAACAGTTACGGATAATATATGTTTACTTGTATG	2900
TGGTACCCTCATTTTGTCTCTCATGGAGGCTTTCAAGCCTTGTTGTAAA	2950
CTGGATAGTTACATATGCTTCCAACAGAAACTAGCATGCAGATTCATATG	3000
CTTTCCTATTCTACTAATTATGTATTGACACACTCGTTGTTTCTTTTGAA	3050
AGATATAAAGTTGGACGAATCGAGCAGCTAGAAACATCTGACCAAGCAAA	3100
AGCCAGAGGTGCTAATACTGTAAGTTTTCTTGGATAGGTCAAGGAGAGTG	3150
TTGCAGACTGTTTTTGATCATTTCTTTTTTTTGTACATTACTTTCATGCTG	3200
TAATTAACTCAATGGCTATTCTGGTCTGATTATCAGATAATTCCAAGGAA	3250
GCTAGTTCAGGTATTAACTCCATCAACAGCAAGCGAGGGAAACATCGGGC	3300
CTGATGCCGTCCATCTTCTTGCTATAAAAGAGGTTTGTTATTTACTTATT	3350
TATCTTATCATGTTCAGTTCATCCAAGTCCTGAAAAATTACACTCTTCTT	3400
TACCAATCTTCCATCAAGCTGTGTAAAGGATTTGGAATTAGAAAATCATT	3450
ATTTGATGCTTTGTTTTATATGCAAGAGGTTCCCTTGAAAAGATCTGTTT	3500
AAGATTCTTTGCACTTGAAAAATTCAATCTTTTTAAGTGAATCCCCTACT	3550
TTCTTACAATGATCATAGTCTGCAATTGCATGTCAAGTAATATCATTCCT	3600
TGTTACTGCATCCCCCTCTTTCTTAATGACCATTGTCTATGTTGTTTTG	3650
TCTCGTGTGCTGGAGAAATGATAGCTGATCCAAGCTGTACATTATCATG	3700
ATTAAGTAGCTGCTCAGGAATTGCCTTTGGTTACATTGCCTAATGGTTTG	3750 3750
ATGTCAATTTTCTTCTGAATCTTTATTTTAGATCAAAATGGAGCTACAA	
AAGTGTTCAACTGTGTATGGATTTGCTTTTGTTGACTGTGCTGCCTTGAG	3800
GTTTTGGGTTGGGTCCATCAGCGATGATGCATCATGTGCTGCTCTTTGGAG	3850
CGTTATTGATGCAGGTAAGCAAGTGTATTCTGTATCTTATGTGTACCATG	3900
TGACTTCCTGTGCATATTTTGGGTTGCAGGAACTAATTCTGAATCACCA	3950
TTTGGTATGTTTTTCCAGGTTTCTCCAAAGGAAGTGTTATATGACAGTA	4000
AAGGTAAACTGCTTGTATCGCCAGTTGTTTTGTTAAACAGAATTTAAGGT	4050
AAATGACACTGGTTAATTTAAAGTGCATACATGTTGAAATATTGCAGGGC	4100
TATCAAGAGAAGCACAAAAGGCTCTAAGGAAATATACGTTGACAGGTACC	4150
ATTTCAGTAGGCAAGCTAACTGACAATTTAACCGCTCACCGAATGATAGG	4200
TCTCTTAAACATTGCTAATGTAGATGATGTTTATGTTTCAATCTAATAGG	4250
GTCTACGGCGGTACAGTTGGCTCCAGTACCACAAGTAATGGGGGATACAG	4300
ATGCTGCTGGAGTTAGAAATATAATAGAATCTAACGGATACTTTAAAGGT	4350
TCTTCTGAATCATGGAACTGTGCTGTTGATGGTCTAAATGAATG	4400
101101GARTON COLUMN COL	

TGCCCTTAGTGCTCTTGGAGAGCTAATTAATCATCTGTCTAGGCTAAAGG	4450
TGTGTTGGCTTGTTTAGTTTTTGCTTTTCACAAATTAAGCAAAGGAACTT	4500
TTCATAACTTACAGTTTCTATCTACTTGCAGCTAGAAGATGTACTTAAGC	4550
ATGGGGATATTTTTCCATACCAAGTTTACAGGGGTTGTCTCAGAATTGAT	4 600
GGCCAGACGATGGTAAATCTTGAGATATTTAACAATAGCTGTGATGGTGG	4650
TCCTTCAGGCAAGTGCATATTTCTTTTTTGATAACTTCAACTAGAGGGCA	4700
GACATAGAAGGAAAAATTCTAATACTTCGTACGGATCTCCAGTAAGTA	4750
AGCCGATTTTTGTTTACCTATGTAGGGACCTTGTACAAATATCTTGATAA	4800
CTGTGTTAGTCCAACTGGTAAGCGACTCTTAAGGAATTGGATCTGCCATC	4850
CACTCAAAGATGTAGAAAGCATCAATAAACGGCTTGATGTAGTTGAAGAA	4900
TTCACGGCAAACTCAGAAAGTATGCAAATCACTGGCCAGTATCTCCACAA	4950
ACTTCCAGACTTAGAAAGACTGCTCGGACGCATCAAGTCTAGCGTTCGAT	5000
CATCAGCCTCTGTGTTGCCTGCTCTTCTGGGGAAAAAAGTGCTGAAACAA	5 050
CGAGTAAGTATCAATCACAAGTTTTCTGAGTAATGCCTTCCATGAGTAGT	5100
ATAGGACTAAAACATTACGGGTCTAGCTAAAGACTGTTCTCCTTCTTTTG	5 150
CAATGTCTGGTTATTCATTACATTTCTCTTAACTTATTGCATTGCAGGTT	5200
AAAGCATTTGGGCAAATTGTGAAAGGGTTCAGAAGTGGAATTGATCTGTT	5250
GTTGGCTCTACAGAAGGAATCAAATATGATGAGTTTGCTTTATAAACTCT	5300
GTAAACTTCCTATATTAGTAGGAAAAAGCGGGCTAGAGTTATTTCTTTC	5350
CAATTCGAAGCAGCCATAGATAGCGACTTTCCAAATTATCAGGTGCCCAT	5400
CTATCTTTCATACTTTACAACAAAATGTCTGTCACTACTCAAAGCAATGC	5450
ATATGGCTTAGATCTCAACTCACACCCCGAGGATCCTAAAGGGATTTGCT	5500
TTTTATTCCTAATGTTTTTGGATGGTTTGATTTATTTCTAACTTGAACTT	5550
ATTAATCTTGTACCAGAACCAAGATGTGACAGATGAAAACGCTGAAACTC	5 600
TCACAATACTTATCGAACTTTTTATCGAAAGAGCAACTCAATGGTCTGAG	5650
GTCATTCACACCATAAGCTGCCTAGATGTCCTGAGATCTTTTGCAATCGC	5700
AGCAAGTCTCTCTGCTGGAAGCATGGCCAGGCCTGTTATTTTTCCCGAAT	5750
CAGAAGCTACAGATCAGAATCAGAAAACAAAAGGGCCAATACTTAAAATC	5800
CAAGGACTATGGCATCCATTTGCAGTTGCAGCCGATGGTCAATTGCCTGT	5850
TCCGAATGATATACTCCTTGGCGAGGCTAGAAGAAGCAGTGGCAGCATTC	5900
ATCCTCGGTCATTGTTACTGACGGGACCAAACATGGGCGGAAAATCAACT	5950
CTTCTTCGTGCAACATGTCTGGCCGTTATCTTTGCCCAAGTTTGTATACT	6000
CGTTAGATAATTACTCTATTCTTTGCAATCAGTTCTTCAACATGAATAAT	6050
AAATTCTGTTTTCTGTCTGCAGCTTGGCTGCTACGTGCCGTGTGAGTCTT	6100
GCGAAATCTCCCTCGTGGATACTATCTTCACAAGGCTTGGCGCATCTGAT	6150
AGAATCATGACAGGAGAGAGTAAGTTTTGTTCTCAAAATACCAATTCCTC	6200
GAACTATTTACTCAGATTTTGTCTGATTGGACAAGGTGGTTTTGCTTTTT	6250
TTTAGGTACCTTTTTGGTAGAATGCACTGAGACAGCGTCAGTTCTTCAGA	6300
ATGCAACTCAGGATTCACTAGTAATCCTTGACGAACTGGGCAGAGGAACT	6350
AGTACTTTCGATGGATACGCCATTGCATACTCGGTAACCTGCTCTTCTCC	6400
TTCAACTTATACTTGTTGATCAACAAAAACATGCAATTCATTTTGCTGAA	6450
ACTTATTGATTTATATCAGGTTTTTCGTCACCTGGTAGAGAAAGTTCAAT	6500
GTCGGATGCTCTTTGCAACACATTACCACCCTCTCACCAAGGAATTCGCG	6550
TCTCACCCACGTGTCACCTCGAAACACATGGCTTGCGCATTCAAATCAAG	6600

Figure 11 (Continued)

ATCTGATTATCAACCACGTGGTTGTGATCAAGACCTAGTGTTCTTGTACC	6650
GTTTAACCGAGGGAGCTTGTCCTGAGAGCTACGGACTTCAAGTGGCACTC	6700
ATGGCTGGAATACCAAACCAAGTGGTTGAAACAGCATCAGGTGCTGCTCA	6750
AGCCATGAAGAGATCAATTGGGGAAAACTTCAAGTCAAG	6900
CTGAGTTCTCAAGTCTGCATGAAGACTGGCTCAAGTCATTGGTGGGTATT	6850
TCTCGAGTCGCCCACAACAATGCCCCCATTGGCGAAGATGACTACGACAC	6900
TTTGTTTTGCTTATGGCATGAGATCAAATCCTCTTACTGTGTTCCCAAAT	6950
AAATGGCTATGACATAACACTATCTGAAGCTCGTTAAGTCTTTTGCTTCT	7000
CTGATGTTTATTCCTCTTAAAAAATGCTTATATATCAAAAAATTGTTTCC	7050
TCGATTATAACAAGATTATATGTATCTGTCGGTTTAGCTATGGTATAT	7100
AATATATGTATGTTCATGAGATTGGTCAAGAGAAATACTCACAAACAGTA	7150
TATTAAGAAGGAAATATGTTTATGCATTAATTTAAGTTTCAAGATAAACT	7200
GCAAATAACCTCGACTAAAGTTGCAAAGACCAAACACAAATTACAAAACT	7250
TATAAGACTTAAGTTCTGAATTCCCTAAAACCAAAAAAAA	7300
TATTTTGTTGCATCTACAAACACAAACCTACATAGTTTATAACTTAC	7350
TCATCACTGAGATTAACATCAGAATCATTCTCCATTTCTTCATCTTCACT	7400
CTCATCATCACCACCACCATGATGATTCTCCTCCTCTTCACGTAACC	7450
TAGCAATCTCACTCTGAGCTCTATCAACAATCTGCTTCTTCTGCAACTCC	7500
AAATCTCTCTGAAAATCAGCTCTCATCTTCTCCAACTCCTTCATTTGCTC	7550
TTTCTTACTCTTCTCCATCTTCTCATAAACCTTCCCAAACCTCTCAACAG	7600
AATCCGCCAACATCTTATACGAAGCAGCGTCATTAACCTTCTTCCTCTCG	7650
TACTCAACCTCATCCTCATCCTCCTCCTCTCTCAGAATCACCAGGACT	7700
ATCCATCATCTCATCAAACCCATTAGACTTATCTAAATAAA	7750
TCATAAACACAAACTCACCTGAATCAACACCACAAGCTAAACCTAAATCC	7800
GACTTGGGCGAAACACAAAGCAACATATCCAACTTATTGAAAAACGACCA	78.50
TTTACTTGAACCTAAACCTGATTTCTCAACCTTAATCTTCTCTTTTCTAT	7900
ACTTCCTCTTCAAGTCATCATTCTCCTACATTGCGTCTCAGATTTC	7950
TCCATCCTTAGCTCCTCACTCTCTCAGCTACTTCATTCCAATCCTC	8000
GTTCCTCAAACTCCTTCTACCCAATTGCAAAAACCTATCTCCCCAAACTT	8050
CAAGCAACACA	8062

Figure 11 (Continued)

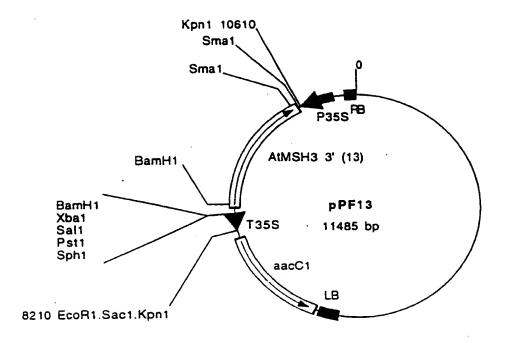


Figure 12

Comments/References: AtMSH3 3' side antisense : AtMSH3 3' (13 = 2104bp) from pUC18/13 Sal1/Sst1/T4 into pCW164 BamH1/T4 in Agrobacterium LBA44O4

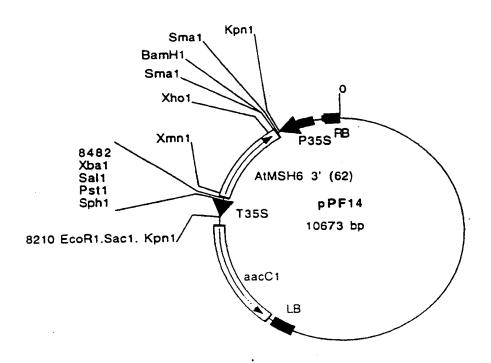


Figure 13

Comments/References: AtMSH6 (S8) 3' side antisens : 62 Sal1/Sst1/T4 (1379bp) into pCW164 BamH1/T4

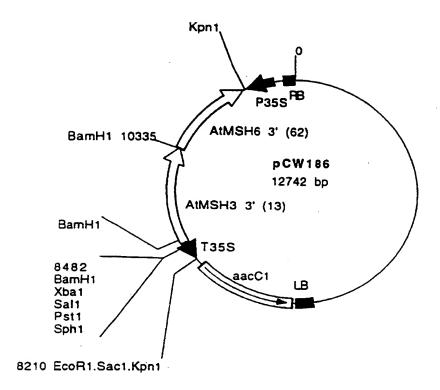


Figure 14

Comments/References: AtMSH6 3'/AtMSH3 3' antisense : AtMSH6 (S8) 3' side (62=1379bp) Sal1/Sst1/T4 into pPF13 (pCW164 AtMSH3 (S5) 3' side (13=2104) antisens)/Sma1. in LBA4404

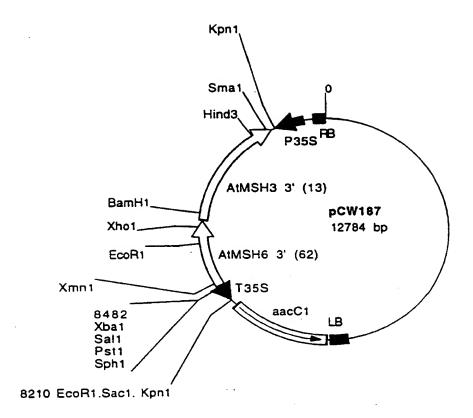


Figure 15

Comments/References: AtMSH3 3'/AtMSH6 3' antisens (D): AtMSH3 (S5) 3' side (13=2104bp) Sal1/Sst1/T4 into pPF14 (AtMSH6 (S8) 3'side (62=1379bp) antisense into pCW164)/Sma1. in LBA4404

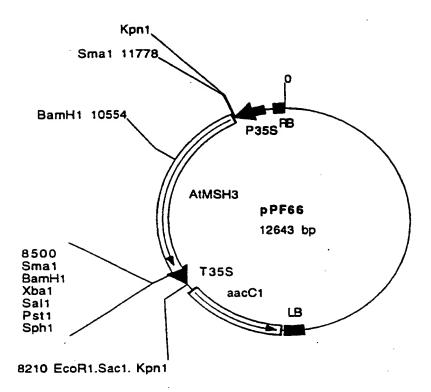


Figure 16

Comments/References: AtMSH3 (S8) complete, sense orientation : pPF26 (3342bp) Sma1 into pCW164 Sma1

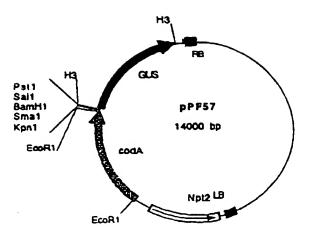


Figure 17

Comments/References: pPZP111 with codA EcoR1 cassette in EcoR1 site and HInd3 GUS cassette in Hind3 site. KanR. All genes under Promoter/terminator 35S

Figure 18

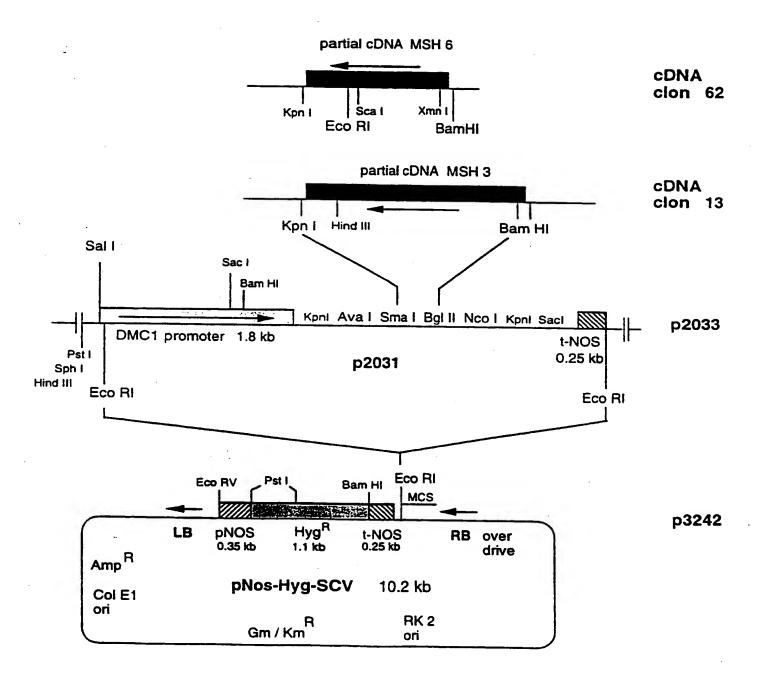
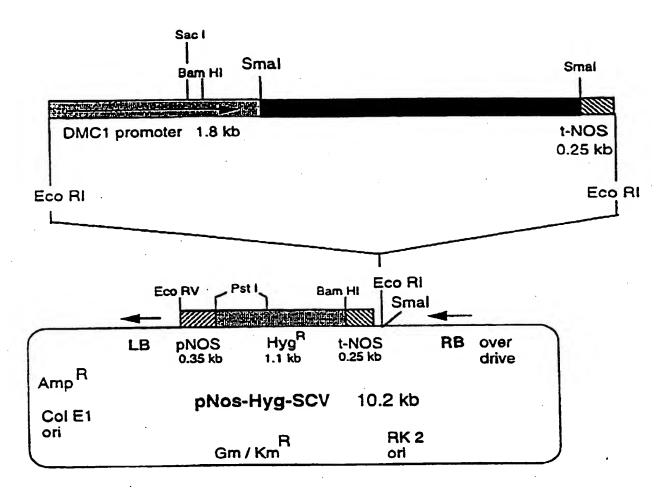


Figure 19

p3243



1

```
SEQUENCE LISTING
```

```
<110>
             Rhone-Poulenc Agro; Betzner, Andreas Stefan; Doutriaux,
             Marie-Pascale; Freyssinet, Georges; Perez, Pascual.
<120>
             Methods for obtaining plant varieties
<130>
             395498C
<150>
             PO9745
<151>
            1997-10-10
<160>
             98
<210>
             1
<211>
             23
<212>
             DNA
<213>
             Artificial sequence
<220>
<221>
            modified_base
<222>
            11
<223>
           · I
<220>
<221>
            modified base
<222>
            14
<223>
             I
<220>
<221>
            modified base
<222>
             17
<223>
<220>
<223>
          Degenerate oligonucleotides UPMU used to isolate AtMSH3 and
             AtMSH6.
<300>
<301>
            Reenan and Kolodner
<302>
            Genetics
<303>
            132
<306>
            963-973
<307>
            1992
<400>
             1
ctggatccac nggnccnaay atg
                                                                      23
<210>
             2
<211>
            23
```

<212>

DNA

2

```
Artificial sequence
<213>
<220>
<221>
              modified base
<222>
              15
<223>
<220>
<221>
              modified_base
<222>
              18
:<223>
 <220>
              Degenerate oligonucleotides DOMU used to isolate AtMSH3 and
 <223>
              AtMSH6.
 <300>
              Reenan and Kolodner
 <301>
 <302>
              Genetics
 <303>
              132
 <306>
              963-973
 < 3.07 >
              1992
 <400>
                                                                       . 23
 ctggatccrt artgngtnrc raa
 <210>
              3
 <211>
              24
 <212>
              DNA
              Artificial sequence
 <213>
 <220>
              MSH3 specific primer 636 for PCR using cDNA of Arabidopsis
 <223>
               thaliana ecotype Columbia
 <400>
               3
                                                                         24
 tgctagtgcc tcttgcaagc tcat
 <210>
               4
 <211>
               27
 <212>
               Artificial sequence
 <213>
 <220>
               Primer AP1 for PCR using cDNA of Arabidopsis thaliana ecotype
 <223>
               Columbia containing adapter sequences ligated to both its
               ends
  <400>
```

.3

ccatcctaat	acgactcact atagggc 27
.210-	5
<210><211>	23
<211>	DNA
<212>	
<213>	Artificial sequence
<220>	•
<223>	Primer AP2 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia containing adapter sequences ligated to both its ends
<400>	5
actcactata	gggctcgagc ggc 23
<210>	6
<211>	-30
<212>	DNA :
<213>	Artificial sequence
\223/	112 02 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
<220>	
<223>	MSH3 specific primer S525 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	6
aggttctgat	tatgtgtgac gctttactta 30
<210>	7
<211>	29
<212>	DNA
<213>	Artificial sequence
<220>	·
<223>	MSH3 specific primer S51 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	7
ggatcgggta	ctgggttttg agtgtgagg 29
<210>	8
<211>	24
<212>	DNA
<213>	Artificial sequence
•	-
<220>	
<223>	MSH3 specific primer 635 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

4

<400>	8
gcacgtgctt	gatggtgttt tcac 24
<210>	9
<211>	28
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	MSH3 specific primer S523 for PCR using cDNA of Arabidopsi thaliana ecotype Columbia
<400>	9
tcagacagta	tccagcatgg cagaagta 28
<210>	10
<211>	33
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	MSH3 specific primer 1S5 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	10
atcccgggat	gggcaagcaa aagcagcaga cga 33
<210>	11
<211>	27
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	MSH3 specific primer S53 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	11
gacaaagagc	gaaatgaggc cccttgg 2°
<210>	12
<211>	1250
<212>	DNA
<213>	Arabidopsis thaliana ecotype Columbia
<223>	Clone 52

<400> 12

cccgggatgg	gcaagcaaaa	gcagcagacg	atttctcgtt	tcttcgctcc	caaacccaaa	60
tccccgactc	acgaaccgaa	tccggtagcc	gaatcatcaa	caccgccacc	gaagatatcc	120
gccactgtat	ccttctctcc	ttccaagcgt	aagcttctct	ccgaccacct	cgccgccgcg	180
tcacccaaaa	agcctaaact	ttctcctcac	actcaaaacc	cagtacccga	tcccaattta	240
caccaaagat	ttctccagag	atttctggaa	ccctcgccgg	aggaatatgt	tcccgaaacg	300
tcatcatcga	ggaaatacac	accattggaa	cagcaagtgg	tggagctaaa	gagcaagtac	360
ccagatgtgg	ttttgatggt	ggaagttggt	tacaggtaca	gattcttcgg	agaagacgcg	420
gagatcgcag	cacgcgtgtt	gggtatttac	gctcatatgg	atcacaattt	catgacggcg	480
agtgtgccaa	catttcgatt	gaatttccat	gtgagaagac	tggtgaatgc	aggatacaag	540
attggtgtag	cgaagcagac	tgaaactgca	gccattaagt	cccatggtgc	aaaccggacc	600
ggcccttttt	cccggggact	greggegrig	tataccaaag	ccacgcttga	agcggctgag	660
gatataagtg	gtggttgtgg	tggtgaagaa	ggttttggtt	cacagagtaa	tttcttggtt	720
tgtgttgtgg	atgagagagt	taagtcggag	acattaggct	gtggtattga	aatgagtttt	780
gatgttagag	teggtgttgt	tggcgttgaa	atttcgacag	gtgaagttgt	ttatgaagag	840
ttcaatgata	atttcatgag	aagtggatta	gaggctgtga	ttttgagctt	gtcaccagct	900
gagctgttgc	ttggccagcc	tctttcacaa	caaactgaga	agtttttggt	ggcacatgct	960
ggacctacct	caaacgttcg	agtggaacgt	gcctcactgg	attgtttcag	caatggtaat	1020
gcagtagatg	aggttatttc	attatgtgaa	aaaatcagcg	caggtaactt	agaagatgat	1080
aaagaaatga	agctggaggc	tgctgaaaaa	ggaatgtctt	gcttgacagt	tcatacaatt	1140
atgaacatgc	cacatctgac	tgttcaagcc	ctcgccctaa	cgttttgcca	tctcaaacag	1200
tttggatttg	aaaggatcct	ttaccaaggg	gcctcatttc	gctctttgtc		1250

<210> 13 <211> 34

<212> DNA

<213> Artificial sequence

<220>

<223> MSH3 specific primer 2S5 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

13 <400 > 34 attoogggto aaaatgaaca agttggtttt agto 14 <210> 27 <211> <212> DNA <213> Artificial sequence <220> MSH3 specific primer S52 for PCR using cDNA of Arabidopsis <223> thaliana ecotype Columbia <400> 14 27 gccacatctg actgttcaag ccctcgc 15 <210> 2110 <211> DNA <212> Arabidopsis thaliana ecotype Columbia <213> Clone 13 <223> <400> 15 gecaeatetg aetgiteaag eedtegeeet aaegtittige eateteaaae agtitiggatt 60 tgaaaggate etttaceaag gggeeteatt tegetetttg teaagtaaca cagagatgae 120 teteteagee aatactetge aacagttgga ggttgtgaaa aataatteag atggategga 180 atotggotoc trattocata atatgaatoa cacacttaca gtatatggtt ccaggottot 240 tagacactgg gtgactcatc ctctatgcga tagaaatttg atatctgctc ggcttgatgc 300 tgtttctgag atttctgctt gcatgggatc tcatagttct tcccagctca gcagtgagtt 360 ggttgaagaa ggttctgaga gagcaattgt atcacctgag ttttatctcg tgctctcctc 420 480 agtettgaca getatgteta gateatetga tatteaaegt ggaataaeaa gaatetttea teggaetget aaageeacag agtteattge agttatggaa getattttae ttgeggggaa 540 gcaaattcag cggcttggca taaagcaaga ctctgaaatg aggagtatgc aatctgcaac 600 tgtgcgatct actottttga gaaaattgat ttctgttatt tcatcccctg ttgtggttga 660 720 caatgccgga aaacttctct ctgccctaaa taaggaagcg gctgttcgag gtgacttgct cgacatacta atcacttcca gcgaccaatt tcctgagctt gctgaagctc gccaagcagt 780 tttagtcatc agggaaaagc tggattcctc gatagcttca tttcgcaaga agctcgctat 840

togaaatttg gaatttotto	aagtgtcggg	gatcacacat	ttgatagagc	tgcccgttga	900
ttccaaggtc cctatgaatt	gggtgaaagt	aaatagcacc	aagaagacta	ttcgatatca	960
tcccccagaa atagtagctg	gcttggatga	gctagctcta	gcaactgaac	atcttgccat	1020
tgtgaaccga-gcttcgtggg	atagtttcct	caagagtttc	agtagatact	acacagattt	1080
taaggctgcc gttcaagctc	ttgctgcact	ggactgtttg	cactcccttt	caactctatc	1140
tagaaacaag aactatgtcc	gtcccgagtt	tgtggatgac	tgtgaaccag	ttgagataaa	1200
catacagtot ggtcgtcatc	ctgtactgga	gactatatta	caagataact	tcgtcccaaa	1260
tgacacaatt ttgcatgcag	aaggggaata	ttgccaaatt	atcaccggac	ctaacatggg	1320
aggaaagagc tgctatatcc	gtcaagttgc	tttaatttcc	ataatggctc	aggttggttc	1380
ctttgtacca gcgtcattcg	ccaagctgca	cgtgcttgat	ggtgttttca	ctcggatggg	1440
tgcttcagac agtatccagc	atggcagaag	tacctttcta	gaagaattaa	gtgaagcgtc	1500
acacataatc agaacctgtt	cttctcgttc	gcttgttata	ttagatgagc	ttggaagagg	1560
cactagcaca cacgacggtg	tagccattgc	ctatgcaaca	ttacagcatc	tcctagcaga	1620
aaagagatgt ttggttcttt	ttgtcacgca	ttaccctgaa	atagctgaga	tcagtaacgg	1680
attcccaggt tctgttggga	cataccatgt	ctcgtatctg	acattgcaga	aggataaagg	1740
cagttatgat catgatgatg	tgacctacct	atataagctt	gtgcgtggtc	tttgcagcag	1800
gagetttggt tttaaggttg	ctcagcttgc	ccagatacct	ccatcatgta	tacgtcgagc	1860
catttcaatg gctgcaaaat	tggaagctga	ggtacgtgca	agagagaa	atacacgcat	1920
gggagaacca gaaggacatg	aagaaccgag	aggcgcagaa	gaatctattt	cggctctagg	1980
tgacttgttt gcagacctga	aatttgctct	ctctgaagag	gacccttgga	aagcattcga	2040
gtttttaaag catgcttgga	agattgctgg	caaaatcaga	ctaaaaccaa	cttgttcatt	2100
ttgacccggg					2110

<210> 16
<211> 29
<212> DNA
<213> Artificial sequence
<220>

<223> MSH3 specific primer S51 for PCR using cDNA of Arabidopsis
thaliana ecotype Columbia

<400> 16

ggatcgggta	ctgggttttg	agtgtgagg							29
<210>	17								
<211>	30								
<212>	DNA								
<213>		al sequence	•						
(213)	ALCILICA	ir bequeene							
<220>									
<223>	_	cific prime ecotype Co			CR usi	ing cDN	A of	Arabidop	sis
<400>	17								
aggttctgat	tatgtgtgac	gctttactta	a						30
<210>	18								
<211>	3522								
<212>	DNA								-
<213>	Arabidop	sis thalia	na eco	type Co	lumbia	a			
<220>									
<221>	CDS								
<222>	(100)	. (3342)							
<223>		ull-length	CDNA	and ded	uced :	sequenc	e of	the enco	oded
	polypept								
<400>	18	· :							
cctaagaaag	cgcgcgaaaa	ttggcaacc	c aagt	tcgcca	tagcc	acgac c	acga	ccttc	60
catttctctt	aaacggagga	gattacgaa	t aaag	gcaatt					99
ato ססכ aa	g caa aag c	ag cag acg	att t	ct cat	ttc t	tc gct	ccc	aaa	147
Mer Gly I.	s Gln Lys G	In Gln Thr	Ile S	Ser Arg	Phe P	he Ala	Pro	Lys	
1	5			10			15	•	
•	- .								
ccc aaa to	cc ccg act c	rac gaa cco	aat	eco ota	acc a	aa tca	tca	aca	195
Pro Lvs Se	er Pro Thr H	dis Glu Pro	Asn	Pro Val	Ala G	lu Ser	Ser	Thr	
FIO Dys 50	20	.10 010 110	25			30			
	. 20								
ככם ככם כי	g aag ata t	cc gcc act	gta 1	tee tte	tct c	ct tcc	aag	cqt	243
Pro Pro P	co Lys Ile S	Ser Ala Thr	. Val	Ser Phe	Ser F	ro Ser	Lys	Arg	
	35	40				45	•	-	
•									
aag ctt c	to too gac	cac ctc qc	acc o	gcg tca	ccc a	aa aaq	cct	aaa	291
	eu Ser Asp H								
Eys Led D	ta ber web t	55			60	-11-		4	
arr rar a	ct cac act o	caa aac co:	a orta	כככ משד	CCC =	at tra	cac	caa	339
	ro His Thr (
65		70		75				80	
		· -							

aga Arg	ttt Phe	ctc Leu	cag Gln	aga Arg 85	ttt Phe	ctg Leu	gaa Glu	ccc Pro	ser 90	ccg Pro	g a g Glu	gaa Glu	tat Tyr	gtt Val 95	ccc Pro	387
gaa Glu	acg Thr	cca Ser	cca Ser 100	tcg Ser	agg Arg	aaa Lys	tac Tyr	aca Thr 105	cca Pro	ttg Leu	gaa Glu	cag Gln	caa Gln 110	gtg Val	gtg Val	435
gag Glu	cta Leu	aag Lys 115	agc Ser	aag Lys	tac Tyr	cca Pro	gat Asp 120	gtg Val	gtt Val	ttg Leu	atg Met	gtg Val 125	gaa Glu	gtt Val	ggt Gly	483
tac Tyr	agg Arg 130	tac Tyr	aga Arg	ttc Phe	ttc Phe	gga Gly 135	gaa Glu	gac Asp	gcg Ala	gag Glu	atc Ile 140	gca Ala	gca Ala	cgc Arg	gtg Val	533
ttg Leu 145	ggt Gly	att Ile	tac Tyr	gct Ala	cat His 150	atg Met	gat Asp	cac His	aat Asn	ttc Phe 155	atg Met	acg Thr	gcg Ala	agt Ser	gtg Val 160	579
cca Pro	aca Thr	ttt Phe	cga Arg	ttg Leu 165	aat Asn	ttc Phe	cat His	gtg Val	aga Arg 170	aga Arg	ctg Leu	gtg Val	aat Asn	gca Ala 175	gga Gly	627
tac Tyr	aag Lys	att Ile	ggt Gly 180	gta Val	gtg Val	aag Lys	cag Gln	act Thr 185	gaa Glu	act Thr	gca Ala	gcc Ala	att Ile 190	aag Lys	tcc Ser	675
cat His	ggt Gly	gca Ala 195	aac Asn	cgg Arg	acc Thr	ggc Gly	cct Pro 200	ttt Phe	ttc Phe	cgg Arg	gga Gly	ctg Leu 205	tcg Ser	gcg Ala	ttg Leu	723
tat Tyr	acc Thr 210	aaa Lys	gcc Ala	acg Thr	ctt Leu	gaa Glu 215	gcg Ala	gct Ala	gag Glu	gat Asp	ata Ile 220	agt Ser	ggt Gly	ggt Gly	tgt Cys	77 1
ggt Gly 225	.ggt Gly	gaa Glu	gaa Glu	ggt Gly	ttt Phe 230	ggt Gly	tca Ser	cag Gln	agt Ser	aat Asn 235	ttc Phe	ttg Leu	gtt Val	tgt Cys	gtt Val 240	819
gtg Val	gat Asp	gag Glu	aga Arg	gtt Val 245	aag Lys	tcg Ser	gag Glu	aca Thr	tta Leu 250	ggc Gly	tgt Cys	ggt Gly	att Ile	gaa Glu 255	atg Met	867
agt Ser	ttt Phe	gat Asp	gtt Val 260	aga Arg	gtc Val	ggt Gly	gtt Val	gtt Val 265	ggc Gly	gtt Val	gaa Glu	att Ile	tcg Ser 270	aca Thr	ggt Gly	915
gaa Glu	gtt Val	gtt Val 275	Tyr	gaa Glu	gag Glu	ttc Phe	aat Asn 280	gat Asp	aat Asn	ttc Phe	atg Met	aga Arg 285	agt Ser	gga Gly	tta Leu	963

												ttg Leu				-	1011
												cat Met				:	1059
												tgt Cys				:	1107
												aaa Lys					1155
												gct Ala 365					1203
gga Gly	atg Met 370	tct Ser	tgc Cys	ttg Leu	aca Thr	gtt Val 375	cat His	aca Thr	att Ile	atg Met	aac Asn 380	atg Met	cca Pro	cat His	ctg Leu		1251
act Thr 385	gtt Val	caa Gln	gcc Ala	ctc Leu	gcc Ala 390	cta Leu	acg Thr	ttt Phe	tgc Cys	cat His 395	ctc Leu	aaa Lys	cag Gln	ttt Phe	gga Gly 400		1299
ttt Phe	gaa Glu	agg Arg	atc Ile	ctt Leu 405	tac Tyr	caa Gln	GJA aaa	gcc Ala	tca Ser 410	Phe	cgc Arg	tct Ser	ttg Leu	tca Ser 415	Ser		1347
aac Asn	aca Thr	gag Glu	atg Met 420	act Thr	ctc Leu	tca Ser	gcc Ala	aat Asn 425	act	ctg Leu	caa Gln	cag Gln	Leu 430	Glu	gtt Val		1395
			Asn					Glu					Phe		aat Asn		1443
		His										ı Leu			tgg Trp	•	1491
gtg Val 465	Thr	cat His	cct Pro	cta Leu	tgc Cys 470	Asp	aga Arg	aat Asn	tt <u>c</u> Le	ata 1 Ile 479	e Ser	gct Ala	cgg Arg	ctt Lev	gat Asp 480		1539
gct Ala	gtt Val	tct Ser	gag Glu	att 11e 489	e Ser	gct Ala	tgc Cys	atg Met	gga Gl _y 490	/ Sei	cat His	agt s Ser	tct Ser	Sei 495	c cag Gln		1587

									• • •							
C:	c age	agt Ser	gag Glu 500	t t g Leu	gct Val	gaa Glu	gaa Glu	ggt Gly 505	tct Ser	gag Glu	aga Arg	gca Ala	att Ile 510	gta Val	tca Ser	1635
Pi	et gag co Gli	g ttt 1 Phe 515	Tyr	ctc Leu	gtg Val	ctc Leu	ser 520	tca Ser	gtc Val	ttg Leu	aca Thr	gct Ala 525	atg Met	tct Ser	aga Arg	1683
Se	a to r Se: 530	: Asp	att Ile	caa Gln	cgt Arg	gga Gly 535	ata Ile	aca	aga Arg	atc Ile	Phe 540	cat His	cgg Arg	act Thr	gct Ala	1731
aa Ly 54	a gco 's Ala 5	aca Thr	gag Glu	ttc Phe	att Ile 550	gca Ala	gtt Val	atg Met	gaa Glu	gct Ala 555	att Ile	tta Leu	ctt Leu	gcg Ala	999 Gly 560	1779
aa Ly	g caa s Gli	a att 1 Ile	cag Gln	cgg Arg 565	ctt Leu	ggc Gly	ata Ile	aag Lys	caa Gln 570	gac Asp	tct Ser	gaa Glu	atg Met	agg Arg 575	agt Ser	1827
at Me	g caa t Gl:	tct Ser	gca Ala 580	act Thr	gtg Val	cga Arg	tct Ser	act Thr 585	ctt Leu	ttg Leu	aga Arg	aaa Lys	ttg Leu 590	att Ile	tct Ser	1875
g t	t att	tca Ser 595	tcc Ser	cct Pro	gtt Val	gtg Val	gtt Val 600	gac Asp	aat Asn	gcc Ala	gga Gly	aaa Lys 605	ctt Leu	ctc Leu	tct Ser	1923
gc Al	c cta a Leu 610	ı Asn	aag Lys	gaa Glu	gcg Ala	gct Ala 615	gtt Val	cga Arg	ggt Gly	gac Asp	ttg Leu 620	ctc Leu	gac Asp	ata Ile	cta Leu	1971
at [] 62	c act e Thi 5	c ccc Ser	agc Ser	gac Asp	caa Gln 630	ttt .Phe	cct Pro	gag Glu	ctt Leu	gct Ala 635	gaa Glu	gct Ala	cgc Arg	caa Gln	gca Ala 640	2019
gt Va	t tta 1 Leu	gtc Val	atc Ile	agg Arg 645	gaa Glu	aag Lys	ctg Leu	gat Asp	tcc Ser 650	tcg Ser	ata	gct Ala	tca Ser	ttt Phe 655	cgc Arg	2067
aa Ly	g aag s Lys	ctc Leu	gct Ala 660	att Ile	cga Arg	aat Asn	ttg Leu	gaa Glu 665	ttt Phe	ctt Leu	caa Gln	gtg Val	tcg Ser 670	GJA aaa	atc Ile	2115
ac Th	a cat r His	ttg Leu 675	ata Ile	gag Glu	ctg Leu	ccc Pro	gtt Val 680	gat Asp	tcc Ser	aag Lys	gtc Val	cct Pro 685	atg His	aat Asn	tgg Trp	2163
gt Va	g aaa 1 Lys 690	Val	aat Asn	agc Ser	acc Thr	aag Lys 695	aag Lys	act Thr	att Ile	cga Arg	tat Tyr 700	cat His	ccc Pro	cca Pro	gaa Glu	2211

	-	-			_	_				_				ctt Leu		2259
			_	_	_		_	_			_	-		agt Ser 735		2307
			_		_	_	_	_		-		_	_	ctg Leu	_	2355
_	_								_		_			gtc Val		2403
				_	_	_	_		_					cag Gln		2451
	_										_			gtc Val		2499
	_			_		_	_		_		-			atc Ile 815		2547
			_					-						gct Ala		2595
			_	_	_	Val								ttc Phe		2643
												Gly		tca Ser		2691
_	Ile	_				_					-		_	gaa Glu		2739
				_		_			_	_		_		tta Leu 895	-	2787
			_			_						_		gcc Ala		2835

WO 99/19492

gca aca Ala Thr															2883
gtc acg Val Thr 930	His														2931
tct gtt Ser Val															2979
ggc agt Gly Ser															3027
ggt ctt Gly Lev															3075
ata cct					Arg					Met				_	3123
gaa gct Glu Ala 1010	Glu			Ala					Thr	_	_				3171
gaa gga Glu Gly 1025			Glu					Glu					Ala		.3219
ggt gad Gly Asi	_	Phe	_		-		Phe	-			_	Glu	_		3267
tgg aaa Trp Lys	. Ala		Glu			Lys					Ile				3315
atc aga		Lys			Cys		Phe	tga	ttta	atc	ttaa	catt	at		3362
agcaact	gca	aggt	cttg	at c	atct	gtta	g tt	gcgt	acta	act	tatg	tgt	atta	gtataa	3422
caagaaa	aaga	gaat	taga	ga g	atgg	attc	t aa	tccg	gtgt	tgc	agta	cat	cttt	tctcca	3482
cccgcat	aaa	aaaa	aaaa	aa a	aaaa	aaaa	a aa	aaaa	aaaa						3522
<210>		19													

<210> 19 <211> 1081 <212> PRT

<213					tide			a ec	ocyp	e co	.Tump	la			
< 4 0 0) >		19												
Met 1	Gly	Lys	Gln	Lys 5	Gln	Gln	Thr	Ile	Ser 10	Arg	Phe	Phe	Ala	Pro 15	Lys
Pro	Lys	Ser	Pro 20	Thr	His	Glu	Pro	Asn 25	Pro	Val	Ala	Glu	Ser 30	Ser	Thr
Pro	Pro	Pro 35	Lys	Ile	Ser	Ala	Thr 40	Val	Ser	Phe	Ser	Pro 45	Ser	Lys	Arg
Lys	Leu 50	Leu	Ser	Asp	His	Leu 55	Ala	Ala	Ala	Ser	Pro 60	Lys	Lys	Pro	Lys
Leu 65	Ser	Pro	His	Thr	Gln 70	Asn	Pro	Val	Pro	Asp 75	Pro	Asn	Leu	His	Gln 80
Arg	Phe	Leu	Gln	Arg 85	Phe	Leu	Glu	Pro	Ser 90	Pro	Glu	Glu	Tyr	Val 95	Pro
Glu	Thr	Ser	Ser 100	Ser	Arg	Lys	Tyr	Thr 105	Pro	Leu	Glu	Gln	Gln 110	Val	Val
Glu	Leu	Lys 115	Ser	Lys	Tyr	Pro	Asp 120	Val	Val	Leu	Met	Val 125	Glu	Val	Gly
Tyr	Arg 130	Tyr	Arg	Phe	Phe	Gly 135	Glu	Asp	Ala	Glu	Ile 140	Ala	Ala	Arg	Val
Leu 145	Gly	Ile	Tyr	Ala	His 150	Met	Asp	His	Asn	Phe 155	Met	Thr	Ala	Ser	Val
Pro	Thr	Phe	Arg	Leu 165	_	Phe	His	Val	Arg 170		Leu	Val	Asn	Ala 175	Gly
Tyr	Lys	Ile	Gly 180	Val	Val	Lys	Gln	Thr 185	Glu	Thr	Ala	Ala	Ile 190	Lys	Sei
His	Gly	Ala 195		Arg	Thr	Gly	Pro 200	Phe	Phe	Arg	Gly	Leu 205	Ser	Ala	Let
Tyr	Thr 210	Lys	Ala	Thr	Leu	Glu 215		Ala	Glu	Asp	Ile 220		Gly	Gly	Суя
Gly 225		Glu	Glu	Gly	Phe 230		Ser	Gln	Ser	235		Leu	. Val	Cys	Va:
Val	Asp	Glu	Arg	Val 245		Ser	Glu	Thr	Lev 250		Cys	Gly	lle	Glu 255	

Ser	Phe	Asp	Val 260	Arg	Val	Gly	Val	Val 265	Gly	Val	Glu	Ile	Ser 270	Thr	Gly
Glu	Val	Val 275	Tyr	Glu	Glu	Phe	Asn 280	Asp	Asn	Phe	Met	Arg 285	Ser	Gly	Leu
Glu	Ala 290	Val	Ile	Leu	Ser	Leu 295	Ser	Pro	Ala	Glu	Leu 300	Leu	Leu	Gly	Gln
305					310					315			Ala	-	320
				325					330	•	_	-	Phe	335	
			340	_		•		345		-		_	11e 350		
•		355		_	_		360					365	Ala		
•	370	•	•			375					380		Pro		
385					390					395			Gln		400
			,	405					410				Leu	415	
			420					425					Leu 430		
	-	435		-			440					445	Phe		
	450					455					460		Arg		
465					470					475					Asp 480
				485					490					495	
			500					505			_		510		Ser
		515	•				520					525			Arg
Ser	Ser	-	Ile	Gln	Arg	Gly		Thr	Arg	Ile	Phe		Arg	Thr	Ala

Ly:	s Ala 5.	Thr	Glu	Phe	Ile 550	Ala	Val	Met	Glu	Ala 555	Ile	Leu	Leu	Ala	Gly 560
Ly	s Gln	Ile	Gln	Arg 565	Leu	Gly	Ile	Lys	Gln 570	Asp	Ser	Glů	Met	Arg 575	Ser
Me	t Gln	Ser	Ala 580	Thr	Val	Arg	Ser	Thr 585	Leu	Leu	Arg	Lys	Leu 590	Ile	Ser
۷a	l Ile	Ser 595	Ser	Pro	Val	Val	Val 600	Asp	Asn	Ala	Gly	Lys 605	Leu	Leu	Ser
Ala	a Leu 610	Asn	Lys	Glu	Ala	Ala 615		Arg	Gly	Asp	Leu 620	Leu	Asp	Ile	Leu
11 62	e Thr	Ser	Ser	Asp	Gln 630	Phe	Pro	Glu	Leu	Ala 635	Glu	Ala	Arg	Gln	Ala 640
۷a	l Leu	Val	Ile	Arg 645	Glu	Lys	Leu	Asp	Ser 650	Ser	Ile	Ala	Ser	Phe 655	Arg
Ly	s Lys	Leu	Ala 660	Ile	Arg	Asn	Leu	Glu 665	Phe	Leu	Gln	Val	Ser 670	Gly	Ile
Th	r His	Leu 675	Ile	Glu	Leu	Pro	Val 680	Asp	Ser	Lys	Val	Pro 685	His	Asn	Trp
۷a	1 Lys 690	Val	Asn	Ser	Thr	Lys 695	Lys	Thr	Ile	Arg	Tyr 700	His	Pro	Pro	Glu
11 70	e Val 5	Ala	Gly	Leu	Asp 710	Glu	Leu	Ala	Leu	Ala 715	Thr	Glu	His	Leu	Ala 720
11	e Val	Asn	Arg	Ala 725	Ser	Trp	ązĄ	Ser	Phe 730		Lys	Ser	Phe	Ser 735	Arg
ту	r Tyr	Thr	Asp 740		Lys	Ala	Ala	Val 745		Ala	Leu	Ala	Ala 750	Leu	Asp
Су	s Leu	His 755	Ser	Leu	Ser	Thr	Leu 760		Arg	Asn	Lys	Asn 765		Val	Arg
Pr	o Glu 770		Val	Asp	Asp	Cys 775	Glu	Pro	Val	Glu	Ile 780	Asn	Ile	Gln	Ser
G1 78	y Arg 5	His	Pro	Val	Leu 790		Thr	Ile	Leu	Gln 795		Asn	Phe	Val	Pro 800
As	n Asp	Thr	Ile	Leu 805		Ala	Glu	Gly	Glu 810		Cys	Gln	Ile	Ile 815	
Gl	y Pro	Asn	Met 820	_	Gly	Lys	Ser	Cys 825	_	: Ile	Àrg	Gln	Val 830		Leu

Il	е	Ser	Ile 835	Met	Ala	Gln	Val	Gly 840	Ser	Phe	Val	Pro	Ala 845	Ser	Phe	Ala
Ly		Leu 850	His	Val	Leu	Asp	Gly 855	Val	Phe	Thr	Arg	Met 860	Gly	Ala	Ser	Asp
Se 86		Ile	Gln	His	Gly	Arg 870	Ser	Thr	Phe	Leu	Glu 875	Glu	Leu	Ser	Ģlu	Ala 880
Se	r	His	Ile	Ile	Arg 885	Thr	Cys	Ser	Ser	Arg 890	Ser	Leu	Val	Ile	Leu 895	Asp
Gl	.u	Leu	Gly	Arg 900	Gly	Thr	Ser	Thr	His 905	Asp	Gly	Val	Ala	Ile 910	Ala	Tyr
Al	la	Thr	Leu 915	Gln	His	Leu	Leu	Ala 920	Glu	Lys	Arg	Cys	Leu 925	Val	Leu	Phe
Vē	al	Thr 930	His	Tyr	Pro	Glu	Ile 935		Glu	Ile	Ser	Asn 940	Gly	Phe	Pro	Gly
	er 45	Val	Gly	Thr	Tyr	His 950	Val	Ser	Tyr	Leu	Thr 955		Gln	Lys	Asp	Lys 960
G:	ly	Ser	Туг	Asp	His 965		Ąsp	Val	Thr	Туг 970		туг	Lys	Leu	Val 975	
G:	ly	Leu	Cys	Ser 980	_	Ser	Phe	Gly	Phe 985		: Val	Ala	Gln	Leu 990		Gln
I	le	Pro	Pro 995		Cys	Ile	Arg	Arg		Ile	e Ser	Met	Ala 1005	Ala	Lys	Leu
G		Ala 1010		. Val	Arg	Ala	Arg		Arg	Asn	n Thr	Arg		Gly	Glu	Pro

Glu Gly His Glu Glu Pro Arg Gly Ala Glu Glu Ser Ile Ser Ala Leu 1040 1025 1030 1035

Gly Asp Leu Phe Ala Asp Leu Lys Phe Ala Leu Ser Glu Glu Asp Pro 1050

Trp Lys Ala Phe Glu Phe Leu Lys His Ala Trp Lys Ile Ala Gly Lys 1065

Ile Arg Leu Lys Pro Thr Cys Ser Phe 1075 1080

<210> 20 <211> 24 <212> DNA Artificial sequence <213>

<220>	
<223>	MSH6 specific primer 638 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	20
tctctaccag c	gtgacgaaaa accg 24
<210>	21
<211>	28
<212>	DNA
<213>	Artificial sequence
(213)	Altiticiai sequence
<220>	
<223>	Primer S81 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	21
cgtcgccttt	agcatecect teetteac 28
<210>	22
<211>	30
<212>	DNA
<213>	Artificial sequence
(213)	MICILICIAL DOGACOOC
<220>	
<223>	MSH6 specific primer S823 for PCR using cDNA of Arabidopsis
(223)	thaliana ecotype Columbia
<400>	22
gcttggcgca	tctaatagaa tcatgacagg 30
<210>	23
<211>	24
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	MSH6 specific primer 637 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	23
gacagcgtca	gttcttcaga atgc 24
.212	24
<210>	24
<211>	33
<212>	DNA

<213>	Artificial sequence	
<220>		
4223>	MSH6 specific primer 1S8 for PCR using cDNA of Arabidopsi thaliana ecotype Columbia	s
<400>	24	
atcccgggat	gcagcgccag agatcgattt tgt	33
<210>	25	
<211>	27	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	MSH6 specific primer S83 for PCR using cDNA of Arabidopsi thaliana ecotype Columbia	s
<400>	25	
cgctatctat	ggctgcttcg aattgag	27
<210>	26	
<211>	1385	
<212>	DNA	
<213>	Arabidopsis thaliana ecotype Columbia	
<223>	Clone 43	
<40C>	26	
cccgggatgc	agegeeagag ategattitg tetttettee aaaaaceeac ggeggegact	60
acgaagggtt	tggtttccgg cgatgctgct agcggcgggg gcggcagcgg aggaccacga	120
tttaatgtga	aggaagggga tgctaaaggc gacgcttctg tacgttttgc tgtttcgaaa	180
tctgtcgatg	aggttagagg aacggatact ccaccggaga aggttccgcg tcgtgtcctg	240
ccgtctggat	ttaagccggc tgaatccgcc ggtgatgctt cgtccctgtt ctccaatatt	300
atgcataagt	ttgtaaaagt cgatgatcga gattgttctg gagagaggag ccgagaagat	360
gttgttccgc	tgaatgattc atctctatgt atgaaggcta atgatgttat tcctcaattt	420
cgttccaata	arggraaaac tcaagaaaga aaccatgctt ttagtttcag tgggagagct	480
gaacttagat	cagtagaaga tataggagta gatggcgatg ttcctggtcc agaaacacca	540
gggatgcgtc	cacgtgcttc tcgcttgaag cgagttctgg aggatgaaat gacttttaag	600
gaggataagg	ttcctgtatt ggactctaac aaaaggctga aaatgctcca ggatccggtr	660

tgtggagaga	agaaagaagt	aaacgaagga	accaaatttg	aatggcttga	gtcttctcga	720
atcagggatg	ccaatagaag	acgtcctgat	gatccccttt	acgatagaaa	gaccttacac	730
ataccacctg	atgttttcaa	gaaaatgtct	gcatcacaaa	agcaatattg	gagtgttaag	840
agtgaatata	tggacattgt	gcttttcttt	aaagtgggga	aattttatga	gctgtatgag	900
ctagatgcgg	aattaggtca	caaggagctt	gactggaaga	tgaccatgag	tggtgtggga	960
aaatgcagac	aggttggtat	ctctgaaagt	gggatagatg	aggcagtgca	aaagctatta	1020
gctcgtggat	ataaagttgg	acgaatcgag	cagctagaaa	catctgacca	agcaaaagcc	1080
agaggtgcta	atactataat	tccaaggaag	ctagttcagg	tattaactcc	atcaacagca	1140
agcgagggaa	acatcgggcc	tgatgccgtc	catcttcttg	ctataaaaga	gatcaaaatg	1200
gagctacaaa	agtgttcaac	tgtgtatgga	tttgcttttg	ttgactgtgc	tgccttgagg	1250
ttttgggttg	ggtccatcag	cgatgatçca	tcatgtgctg	ctcttggagc	gttattgatg	1320
caggiticic	caaaggaagt	gttatatgac	agtaaagggc	tatcaagaga	agcacaaaag	1330
gctctaagga	aatatacgtt	gacagggtct	acggcggtac	agttggctcc	agtaccacaa	1440
gtaatggggg	atacagatgc	tgctggagtt	agaaatataa	tagaatctaa	cggatacttt	1500
aaaggttctt	ctgaatcatg	gaactgtgct	gttgatggtc	taaatgaatg	tgatgttgcc	1560
cttagtgctc	ttggagagct	aattaatcat	ctgtctaggc	taaagctaga	agatgtactt	1620
aagcatgggg	atatttttcc	ataccaagtt	tacaggggtt	gtctcagaat	tgatggccag	1680
acgatggtaa	atcttgagat	atttaacaat	agctgtgatg	gtggtccttc	agggaccttg	1740
tacaaatatc	ttgataactg	tgttagtcqa	actggtaagc	gactcttaag	gaattggatc	1800
tgccatccac	tcaaagatgt	agaaagcatc	aataaacggc	ttgatgtagt	tgaagaattc.	1860
acggcaaact	cagaaagtat	gcaaatcact	ggccagtato	tccacaaact	tccagactta	1920
gaaagactgc	tcggacgcat	caagtctago	gttcgatcat	cagcctctgt	gttgcctgct	1980
cttctgggga	aaaagtgct	gaaacaacga	gttaaagcat	ttgggcaaat	tgtgaaaggg	2040
ttcagaagtg	gaattgatct	gttgttggct	ctacagaagg	aatcaaatat	gatgagtttg	2100
ctttataaac	tctgtaaact	tcctatatta	gtaggaaaaa	gegggetage	a gttatttctt	2160
tctcaattcg	aagcagccat	agatagcg				2188

<210> <211> <212> <213> <223> <400>	27 1385 DNA Arabidop: Clone 62	sis thalian	a ecotype Co	olumbia	·	
catcageete	tgtgttgcct	gctcttctgg	ggaaaaaagt	gctgaaacaa	cgagttaaag	60
catttgggca	aattgtgaaa	gggttcagaa	gtggaattga	tctgttgttg	gctctacaga	120
aggaatcaaa	tatgatgagt	ttgctttata	aactctgtaa	acttcctata	ttagtaggaa	180
aaagcgggct	agagttattt	ctttctcaat	tcgaagcagc	catagatagc	gactttccaa	240
attatcagaa	ccaagatgtg	acagatgaaa	acgctgaaac	tctcacaata	cttatcgaac	300
tttttatcga	aagagcaact	caatggtctg	aggtcattca	caccataagc	tgcctagatg	360
tcctgagatc	ttttgcaatc	gcagcaagtc	tctctgctgg	aagcatggcc	aggcctgtta	420
tttttcccga	atcagaagct	acagatcaga	atcagaaaac	aaaagggcca	atacttaaaa	480
tccaaggact	atggcatcca	tttgcagttg	cageegatgg	tcaattgcct	gttccgaatg	540
atatactcct	tggcgaggct	agaagaagca	gtggcagcat	tcatcctcgg	tcattgttac	600
tgacgggacc	aaacatgggc	ggaaaatcaa	ctcttcttcg	tgcaacatgt	ctggccgtta	660
tctttgccca	acttggctgc	tacgtgccgt	gtgagtcttg	cgaaatctcc	ctcgtggata	720
CTATCTTCAC	aaggettgge	gcatctgata	gaatcatgac	aggagagagt	acctttttgg	780
tagaatgcac	tgagacagcg	tcagttcttc	agaatgcaac	tcaggattca	ctagtaatcc	840
ttgacgaact	gggcagagga	actagtactt	tcgatggata	cgccattgca	tactcggttt	900
ttcgtcacct	ggtagagaaa	gttcaatgtc	ggatgctctt	tgcaacacat	taccaccctc	960
tcaccaagga	attegegtet	cacccacgtg	tcacctcgaa	acacatggct	tgcgcattca	1020
aatcaagatc	tgattatcaa	ccacgtggtt	gtgatcaaga	cctagtgttc	ttgtaccgtt	1080
taaccgaggg	agcttgtcct	gagagctacg	gacttcaagt	ggcactcatg	gctggaatac	1140
caaaccaagt	ggttgaaaca	gcatcaggtg	ctgctcaagc	catgaagaga	tcaattgggg	1200
aaaacttcaa	gtcaagtgag	ctaagatctg	agttctcaag	tctgcatgaa	gactggctca	1260
agtcattggt	gggtatttct	cgagtcgccc	acaacaatgc	ccccattggc	gaagatgact	1320
acgacacttt	gttttgctta	tggcatgaga	tcaaatcctc	ttactgtgtt	cccaaataac	1380

ccggg	1385
<210>	28
<211>	34
<212>	DNA
<213>	Artificial sequence
1227	
<220>	
<223>	MSH6 specific primer 2S8 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	26
atcccgggtt	atttgggaac acagtaagag gatt 34
<210>	29
<211>	27
<212>	DNA
<213>	Artificial sequence
<220>	·
<223>	MSH6 specific primer S82 for PCR using cDNA of Arabidopsis
(223)	thaliana ecotype Columbia
<400>	29
gcgttcgatc	atcagcctct gtgttgc 27
<210>	30
<211>	3606
<212>	DNA
<213>	Arabidopsis thaliana ecotype Columbia
	••
<220>	•
<221>	CDS
<222>	(142)(3468)
<223>	AtMSH6 full-length cDNA and deduced sequence of the encoded
	polypeptide
<400>	30
aaaagttgag	ccctgaggag tatcgtttcc gccatttcta cgacgcaagg cgaaaatttt 60
tggcgccaat	ctttcccccc tttcgaattc tctcagctca aaacatcgtt tctctctcac 120
teteteteae	aattccaaaa a atg cag cgc cag aga tcg att ttg tct ttc 171 Met Gln Arg Gln Arg Ser Ile Leu Ser Phe 1 5 10

	aaa Lys										219
	agc Ser										267
	gat Asp 45										315
	gat Asp										363
	gtc. Val										411
	tcc Ser										459
	gat Asp			-	_	-	-	_	_	_	507
	tca Ser 125										555
	aat Asn										603
	aga Arg										651
	cct Pro										699
	cga Arg										747
	ttg Leu 205										795

cys Cys	gga Gly 220	gag Glu	aag Lys	aaa Lys	Glu	gta Val 225	aac Asn	gaa Glu	gga Gly	acc Thr	aaa Lys 230	ttt Phe	gaa Glu	tgg Trp	ctt Leu	843
gag Glu 235	tct Ser	cct Ser	cga Arg	acc Ile	agg Arg 240	gat Asp	gcc Ala	aat Asn	aga. Arg	aga Arg 245	cgt Arg	cct Pro	gat Asp	gat Asp	ccc Pro 250	891
ctt Leu	tac Tyr	gat Asp	aga Arg	aag Lys 255	acc Thr	tta Leu	cac His	ata Ile	cca Pro 260	cct Pro	gat Asp	gtt Val	ttc Phe	aag Lys 265	aaa Lys	939
atg Met	tct Ser	gca Ala	tca Ser 270	caa Gln	aag Lys	caa Gln	tat Tyr	tgg Trp 275	agt Ser	gtt Val	aag Lys	agt Ser	gaa Glu 280	tat Tyr	atg Met	987
gac Asp	att Ile	gtg Val 295	ctt Leu	ttc Phe	ttt Phe	aaa Lys	gtg Val 290	Gly ggg	aaa Lys	ttt Phe	tat Tyr	gag Glu 295	ctg Leu	tat Tyr	gag Glu	1035
cta Leu	gat Asp 300	gcg Ala	gaa Glu	tta Leu	ggt Gly	cac His 305	aag Lys	gag Glu	ctt Leu	gac Asp	tgg Trp 310	aag Lys	atg Met	acc Thr	atg Met	1083
agt Ser 315	Gly	gtg Val	gga Gly	aaa Lys	tgc Cys 320	aga Arg	cag Gln	gtt Val	ggt Gly	atc Ile 325	Ser	gaa Glu	agt Ser	GJA aaa	ata Ile 330	1131
gat Asp	gag Glu	gca Ala	gtg Val	caa Gln 335	aag Lys	cta Leu	tta Leu	gct Ala	cgt Arg 340	Gly	tat Tyr	aaa Lys	gtt Val	gga Gly 345	cga Arg	1179
ato	gag Glu	cag Gln	cta Leu 350	Glu	aca Thr	tct Ser	gac	caa Gln 355	Ala	aaa Lys	gcc Ala	aga Arg	ggt Gly 360	Ala	aat Asn	1227
act Thr	ata :Ile	att Ile 365	Pro	agg Arg	aag Lys	cta Leu	.gtt Val 370	Gln	gta Val	tta Leu	act Thr	cca Pro 375	Ser	aca Thr	gca Ala	1275
ago Sei	gag Glu 380	ı Gly	aac ⁄Asn	atc Ile	ggg Gly	Pro 385) Asp	gco Ala	gto Val	cat His	CET S Lev 39(ı Lev	gct Ala	ata Ile	a aaa e Lys	1323
gag Glu 399	ı Ile	aaa Lys	atg Met	gag Glu	cta Leu 400	ı Glr	a aag a Lys	g tgt s Cys	tca Sei	a act	r Vai	g tat l Tyr	gga Gly	ttt Phe	gct Ala 410	1371
tt: Pho	t gtt e Val	gad L Asi	c tgt o Cys	gct Ala 415	a Ala	tto a Le	g ago	g tti	t tgg Trj 42	o Va	t ggg	g tco y Sei	c ato	age Se: 42	c gat r Asp 5	1419

gat Asp	gca Ala	tca Ser	tgt Cys 430	gct Ala	gct Ala	ctt Leu	gga Gly	gcg Ala 435	tta Leu	t t g Leu	atg Met	cag Gln	gtt Val 440	tct Ser	cca Pro	1467
					gac Asp											1515
gct Ala	cta Leu 460	agg Arg	aaa Lys	tat Tyr	acg Thr	ttg Leu 465	aca Thr	Gly	tct Ser	acg Thr	gcg Ala 470	gta Val	cag Gln	ttg Leu	gct Ala	1563
					atg Met 480											1611
					gga Gly											1659
					cta Leu									Ala		1707
gga Gly	gag Glu	cta Leu 525	att Ile	aat Asn	cat His	ctg Leu	tct Ser 530	agg Arg	cta Leu	aag Lys	cta Leu	gaa Glu 535	Asp	gta Val	ctt Leu	1755
					ttt Phe							Gly				1803
att Ile 555	Asp	ggc Gly	cag Gln	acg Thr	atg Met 560	Val	aat Asn	ctt Leu	gag Glu	ata Ile 565	Phe	aac Asn	aat Asn	ago Ser	tgt Cys 570	1851
gat Asp	ggt Gly	ggt Gly	cct Pro	tca Ser 575	gly	acc Thr	ttg Leu	tac Tyr	aaa Lys 580	Tyr	ctt Leu	gat Asp	aac Asn	tgt Cys 585	: Val	1899
				Lys					Asn					Pro	ctc Leu	1947
aaa Lys	gat Asp	gta Val	Glu	ago Ser	ato	aat Asn	aaa Lys 610	Arg	r ctt	gat Asp	gta Val	a gtt l Val 615	lGli	ı gaz ı Glı	a ttc 1 Phe	1995
		Asr.					Glr					n Tyi			c aaa s Lys	2043

ctt Leu 635											aag Lys			_		2091
											aaa Lys					2139
	_	_		_							Gly					2187
	_	_	_	_	-		_	_	_		aat Asn					2235
											gga Gly 710					2283
											gat Asp					2331
											gct Ala					2379
Ile	Leu	Ile	Glu 750	Leu	Phe	Ile	Glu	Arg 755	Ala	Thr	caa Gln	Trp	Ser 760	Glu	Val	2427
Ile	His	Thr 765	Ile	Ser	Cys	Leu	Asp 770	Val	Leu	Arg	Ser	Phe 775	Ala	Ile	gca Ala	2475
Ala	Ser 780	Leu	Ser	Ala	Gly	Ser 785	Met	Ala	Arg	Pro	Val 790	Ile	Phe	Pro	gaa Glu	2523
						Asn					Gly				aaa Lys 810	2571
					His					Ala					ttg Leu	2619
				Asp					Glu					Ser	ggc Gly	2667

_			cct Pro			-	_	_				_				2715
			ctt Leu													2763
			tac Tyr													2811
			aca Thr													2859
			ttg Leu 910													2907
			gat Asp													2955
_			gat Asp				-		_	-		_		_		3003
			gtt Val						_							3051
		_	gaa Glu					_	_		_			_	•	3099
			ttc Phe 990								Arg					3147
	Asp		gtg Val			Tyr				Gly						3195
Ser			ctt Leu		Val				Gly							3243
	Glu		gca Ala	Ser				Ala					Ile			3291

Glu Asn Phe Lys Ser Ser Glu Leu Arg Ser Glu Phe Ser Ser Leu His 1055 1060 1065	3339
gaa gac tgg ctc aag tca ttg gtg ggt att tct cga gtc gcc cac aac Glu Asp Trp Leu Lys Ser Leu Val Gly Ile Ser Arg Val Ala His Asn 1070 1075 1080	3387
aat gcc ccc att ggc gaa gat gac tac gac act ttg ttt tgc tta tgg Asn Ala Pro Ile Gly Glu Asp Asp Tyr Asp Thr Leu Phe Cys Leu Trp 1085 1090 1095	3435
cat gag atc aaa tcc tct tac tgt gtt ccc aaa taaatggcta His Glu Ile Lys Ser Ser Tyr Cys Val Pro Lys 1100 1105	3478
tgacataaca ctatctgaag ctcgttaagt cttttgcctc tctgatgttt attcctctt	a 3538
aaaaatgott atatatoaaa aaattgttto otogattaaa aaaaaaaaa aaaaaaaaa	a 3598
aaaaaaaa	3606
<pre><210></pre>	
<400> 31	
<pre><400> 31 Met Gln Arg Gln Arg Ser Ile Leu Ser Phe Phe Gln Lys Pro Thr Ala 1 5 10 15</pre>	
Met Gln Arg Gln Arg Ser Ile Leu Ser Phe Phe Gln Lys Pro Thr Ala	
Met Gln Arg Gln Arg Ser Ile Leu Ser Phe Phe Gln Lys Pro Thr Ala 1 5 2 10 2 15 2 15 Ala Ala Ser Gly Gly Gly	
Met Gln Arg Gln Arg Ser Ile Leu Ser Phe Phe Gln Lys Pro Thr Ala 15 Ala Thr Thr Lys Gly Leu Val Ser Gly Asp Ala Ala Ser Gly Gly Gly 25 Gly Ser Gly Gly Pro Arg Phe Asn Val Arg Glu Gly Asp Ala Lys Gly	
Met Gln Arg Gln Arg Ser Ile Leu Ser Phe Phe Gln Lys Pro Thr Ala 15 Ala Thr Thr Lys Gly Leu Val Ser Gly Asp Ala Ala Ser Gly Gly Gly Gly 25 Asp Ala Ser Gly Gly Gly Gly 30 Asp Ala Ser Val Arg Phe Ala Val Ser Lys Ser Val Asp Glu Val Arg	
Met Gln Arg Gln Arg Ser Ile Leu Ser Phe Phe Gln Lys Pro Thr Ala 15 Thr Thr Lys Gly Leu Val Ser Gly Asp Ala Ala Ser Gly Gly Gly Gly Ser Gly Gly Gly Asp Ala Arg Phe Asn Val Arg Glu Gly Asp Ala Lys Gly Asp Ala Ser Val Arg Phe Ala Val Ser Lys Ser Val Asp Glu Val Arg Gly Thr Asp Thr Pro Pro Glu Lys Val Pro Arg Arg Val Leu Pro Ser	
Met Gln Arg Gln Arg Ser Ile Leu Ser Phe Phe Gln Lys Pro Thr Ala Ala Thr Thr Lys Gly Leu Val Ser Gly Asp Ala Ala Ser Gly Gly Gly Gly Ser Gly Gly Pro Arg Phe Asn Val Arg Glu Gly Asp Ala Lys Gly Asp Ala Ser Val Arg Phe Ala Val Ser Lys Ser Val Asp Glu Val Arg Gly Thr Asp Thr Pro Pro Glu Lys Val Pro Arg Arg Val Leu Pro Ser 65 Phe Phe Gln Lys Pro Ala Glu Ser Ala Gly Asp Ala Ser Ser Leu Phe Ser	

Мес	Lys 130	Ala	Asn	Asp	Val	Ile 135	Pro	Gln	Phe	Arg	Ser 140	Asn	Asn	Gly	Lys
Thr 145	Gln	Glu	Arg	Asn	His 150	Ala	Phe	Ser	Phe	Ser 155	Gly	Arg	Ala	Glu	Leu 160
Arg	Ser	Val	Glu	Asp 165	Ile	Gly	Val	Asp	Gly 170	Asp	Val	Pro	Gly	Pro 175	Glü
Thr	Pro	Gly	Met 180	Arg	Pro	Arg	Ala	Ser 185	Arg	Leu	Lys	Arg	Val 190	Leu	Glu
Asp	Glu	Met 195	Thr	Phe	Lys	Glu	Asp 200	Lys	Val	Pro	Val	Leu 205	Asp	Ser	Asn
Lys	Arg 210	Leu	Lys	Met	Leu	Gln 215	Asp	Pro	Val	Cys	Gly 220	Glu	Lys	Lys	Glu
Val 225	Asn	Glu	Gly	Thr	Lys 230	Phe	Glu	Trp.	Leu	Glu 235	Ser	Ser	Arg	Ile	Arg 240
Ąsp	Ala	Asn	Arg	Arg 245	Arg	Pro	Asp	Asp	Pro 250	Leu	Tyr	Asp	Arg	Lys 255	Thr
Leu	His	Ile	Pro 260	Pro	Asp	Val	Phe	Lys 265	Lys	Met	Ser	Ala	Ser 270	Gln	Lys
Gln	Tyr	Trp 275	Ser	Val	Lys	Ser	Glu 280	Tyr	Met	Asp	Ile	Val 285	Leu	Phe	Phe
	290	_				295					Asp 300				
His 305	Lys	Glu	Leu	-	Trp 310	Lys	Met	Thr	Met	Ser 315	Gly	Val	Gly	Lys	Cys 320
Arg	Gln	Val	Gly	Ile 325	Ser	Glu	Ser	Gly	Ile 330	_	Glu	Ala	Val	Gln 335	
Leu	Leu	Ala	Arg 340	Gly	Tyr	Lys	Val	Gly 345	Arg	Ile	Glu	Gln	Leu 350	Glu	Thr
Ser	Asp	Gln 355	Ala	Lys	Ala	Arg	Gly 360	Ala	Asn	Thr	Ile	11e 365	Pro	Arg	Lys
Leu	Val 370	Gln	Val	Leu	Thr	Pro 375	Ser	Thr	Ala	Ser	Glu 380	Gly	Asn	Ile	Gly
Pro 385	Asp	Ala	Val	His	Leu 390	Leu	Ala	Ile	Lys	Glu 395	Ile	Lys	Met	Glu	Leu 400
Gln	Lys	Cys	Ser	Thr 405	Val	Tyr	Gly	Phe	Ala 410		Val	Asp	Cys	Ala 415	Ala

Leu	Arg	Phe	Trp 420	Val	Gly	Ser	Ile	Ser 425	Asp	Asp	Ala	Ser	Cys 430	Ala	Ala
Leu	Gly	Ala 435	Leu	Leu	Met	Gln	Val 440	Ser	Pro	Lys	Glu	Val 445	Leu	Tyr	Asp
Ser	Lys 450	Gly	Leu	Ser	Arg	Glu 455	Ala	Gln	Lys	Ala	Leu 460	Arg	Lys	Tyr	Thr
Leu 465	Thr	Gly	Ser	Thr	Ala 470	Val	Gln	Leu	Ala	Pro 475	Val	Pro	Gln	Val	Met 480
Gly	Asp	Thr	Asp	Ala 485	Ala	Gly	Val	Arg	Asn 490	Ile	Ile	Glu	Ser	Asn 495	Gly
Tyr	Phe	Lys	Gly 500	Ser	Ser	Glu	Ser	Trp 505	Asn	Cys	Ala	Val	Asp 510	Gly	Leu
Asn	Glu	Cys 515	Asp	Val	Ala	Leu	Ser 520	Ala	Leu	Gly	Glu	Leu 525	Ile	Asn	His
Leu	Ser 530	Arg	Leu	Lys	Leu	Glu 535	Asp	Val	Leu	Lys	His 540	Gly	Asp	Ile	Phe
Pro 545	Tyr	Gln	Val	Tyr	Arg 550	Gly	Cys	Leu	Arg	Ile 555	Asp	Gly	Gln	Thr	Met 560
Val	Asn	Leu	Glu	Ile 565	Phe	Asn	Asn	Ser	Cys 570	Asp	Gly	Gly	Pro	Ser 575	Gly
Thr	Leu	Tyr	Lys 580	Tyr	Leu	Asp	Asn	Cys 585	Val	Ser	Pro	Thr	Gly 590	Lys	Arg
Leu	Leu	Arg 595	Asn	Trp	Ile	Cys	His 600	Pro	Leu	Lys	Asp	Val 605	Glu	Ser	Ile
Asn	Lys 610	Arg	Leu	Asp	Val	Val 615	Glu	Glu	Phe	Thr	Ala 620	Asn	Ser	Glu	Ser
Met 625	Gln	Ile	Thr	Gly	Gln 630	Tyr	Leu	His	Lys	Leu 635	Pro	Asp	Leu	Glu	Arg 640
Leu	Leu	Gly	Arg	Ile 645	Lys	Ser	Ser	Val	Arg 650	Ser	Ser	Ala	Ser	Val 655	Leu
Pro	Ala	Leu	Leu 660	Gly	Lys	Lys	Val	Leu 665	Lys	Gln	Arg	Val	Lys 670	Ala	Phe
Gly	Gln	Ile 675	Val	Lys	Gly	Phe	Arg 680	Ser	Gly	Ile	Asp	Leu 685	Leu	Leu	Ala
Leu	Gln 690	Lys	Glu	Ser	Asn	Met 695	Met	Ser	Leu	Leu	Tyr	Lys	Leu	Cys	Lys

705	710	110	Dea	Val	710	Lys	3-1	Gly	Dea	715	Dea	FIIC	Leu	361	720
Phe	Glu	Ala	Ala	Ile 725	Asp	Ser	Asp	Phe	Pro 730	Asn	Tyr	Gln	Asn	Gln 735	Asp
Val	Thr	qaA	Glu 740	Asn	Ala	Glu	Thr	Leu 745	Thr	Ile	Leu	Ile	Glu 750	Leu	Phe
Ile	Glu	Arg 755	Ala	Thr	Gln	Trp	Ser 760	Glu	Val	Ile	His	Thr 765	Ile	Ser	Cys
Leu	Asp 770	Val	Leu	Arg	Ser-	Phe 775	Ala	Ile	Ala	Ala	Ser 780	Leu	Ser	Ala	Gly
Ser 785	Met	Ala	Arg	Pro	Val 790	Ile	Phe	Pro	Glu	Ser 795	Glu	Ala	Thr	Asp	Gln 800
Asn	Gln	Lys	Thr	Lys 805	Gly	Pro	Ile	Leu	Lys 810	Ile	Gln	Gly	Leu	Trp 815	His
Pro	Phe	Ala	Val 820	Ala	Ala	Asp	Gly	Gln 825	Leu	Pro	Val	Pro	Asn 830	Asp	Ile
Leu	Leu	Gly 835	Glu	Ala	Arg	Arg	Ser 840	Ser	Gly	Ser	Ile	His 845	Pro	Arg	Ser
Leu	Leu 850	Leu	Thr	Gly	Pro	Asn 855	Met	Ģly	Gly	Lys	Ser 860	Thr	Leu	Leu	Arg
Ala 865	Thr	Cys	Leu	Ala	Val 870	Ile	Phe	Ala	Gln	Leu 875	Gly	Cys	Tyr	Val	880
Cys	Glu	Ser	Суѕ	Glu 885		Ser	Leu	Val	Asp 890	Thr	Ile	Phe	Thr	Arg 895	Lev
Gly	Ala	Ser	Asp 900	Arg	Ile	Met	Thr	Gly 905	Glu	Ser	Thr	Phe	Leu 910	Val	Glu
Cys	Thr	Glu 915	Thr	Ala	Ser	Val	Leu 920	Gln	Asn	Ala	Thr	Gln 925	Asp	Ser	Lev
Val	Ile 930	Leu	Asp	Glu	Leu	Gly 935	Arg	Gly	Thr	Ser	Thr 940	Phe	Asp	Gly	Туп
Ala 945	Ile	Ala	Tyr	Ser	Val 950	Phe	Arg	His	Leu	Val 955	Glu	Lys	Val	Gln	Суs 960
Arg	Met	Leu	Phe	Ala 965	Thr	His	Tyr	His	Pro 970	Leu	Thr	Lys	Glu	Phe 975	
Ser	His	Pro	Arg 980	Val	Thr	Ser	Lys	His 985		Ala	Cys	Ala	Phe 990	-	Sei

Arg Ser	Asp Tyr	r Gln	Pro Ar	g Gly 1000	Cys	Asp	Gln	Asp 1	Leu L005	Val	Pne	Leu			
Tyr Arg 1010	Leu Thi	r Glu	Gly Al 101		Pro	Glu	Ser	Tyr 1020	Gly	Leu	Gln	Val			
Ala Leu 1025	Met Ala		Ile Pr	o Asn	Gln	Val	Val 1035	Glu	Thr	Ala	Ser 10	Gly 040		-	
Ala Ala	Gln Ala	a Met 1045	Lys Ar	g Ser	Ile	Gly 1050	Glu	Asn	Phe	Lys	Ser 1055	Ser			
Glu Leu	Arg Se 106		Phe Se	r Ser	Leu 1065	His	Glu	Asp	Trp	Leu 1070	Lys	Ser			
Leu Val	Gly Il 1075	e Ser	Arg Va	1080		Asn	Asn	Ala	Pro 1085	Ile	Gly	Glu			
Asp Asp 1090		p Thr	Leu Pi		: Leu	Trp	His	Glu 1100	Ile	Lys	Ser	Ser			
Tyr Cys 1105	Val Pr	o Lys													
			,												
<210>	3	2													
<211>	2	24													
<212>		NA												-	
<213>	А	Artifi	cial s	equenc	ce										•
<220>															
<223>			d prim atelli		c PCR	am <u>r</u>	olifi	.cati	on c	of Al	HGEN	IEA			
<400>	. 3	32	-					·						_	
accatgo	ata gc1	ttaaac	tt ctt	g									-	24	Į.
<210>	. :	33									*				
<211>		22										•			
<212>		DNA													
<213>	j	Artifi	cial s	equen	ce										
<220>	•														
<223>			se prim satelli		r PC	R am	plif	icat	ion (of A'	THGE	NEA			
<400>		33						•							
acataa	ccac aa	atagg	ggt gc											2	2

WO 99/19492 PCT/EP98/06977

<210>	34				
<211>	18				
<212>	DNA				
<213>	Artificial sequence				
<220>					
<223>	Forward primer DMCIN-A for PCR thaliana ssp. Landsberg erecta		DNA	of	Arabidopsis
<400>	34				
gaagcgatat	tgttcgtg				18
<210>	35				
<211>	18				
<212>	DNA				
<213>	Artificial sequence				
<220>	·				
<223>	Reverse primer DMCIN-B for PCR thaliana ssp. Landsberg erecta		DNA	of	Arabidopsis
<400>	35				
agattgcgag	aacattcc				18
<210>	36	•			
<211>	31				
<212>	DNA				
<213>	Artificial sequence	•			
(213)	Arcificial sequence				
<220>					
<223>	Forward primer DMCIN-1 for PCR thaliana ssp. Landsberg erecta		DNA	of	Arabidopsis
<400>	36		-		
acgcgtcgac	tcagctatga gattactcgt g				31
<210>	37				
<211>	29				
<212>	DNA				
<213>					
(213)	Artificial sequence				
<220>					
<223>	Reverse primer DMCIN-2 for PCR thaliana ssp. Landsberg erecta		DNA	of	Arabidopsis
<400>	37				
gctctagatr	teregeteta agaereret				29

<210>	38
<211>	32
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer DMCIN-3 for PCR on genomic DNA of Arabidopsis thaliana ssp. Landsberg erecta "Ler"
<400>	38
gctctagaġc t	tctcttaag taagtgattg at 32
<210>	39
<211>	48
<212>	DNA
<213>	Artificial sequence
<220>	Designed and DMCTN A few DCD and appearing DNN of North Administra
<223>	Reverse primer DMCIN-4 for PCR on genomic DNA of Arabidopsis thaliana ssp. Landsberg erecta "Ler"
<400>	39
tecceeggge t	cgagagate tecatggttt etteagetet atgaatee 48
<210>	40
<211>	26
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer DMCla for PCR on genomic DNA of Arabidopsis thaliana ssp. Landsberg erecta "Ler"
<400>	40
acgcgtcgac g	gaattegeaa gtgggg 26
<210>	41
<211>	38
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Payerse primer DMC1b for DCD on denomic DNA of Archidonsis
\263 >	Reverse primer DMClb for PCR on genomic DNA of Arabidopsis thaliana ssp. Landsberg erecta "Ler"
<400>	41

tccatggaga	tetecegggt accgattige tiegaggg	38
·		
<210>	42	
<211>	20	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of ATEAT1 SSLP market Arabidopsis thaliana subspecies	r ir
<400>	42	
gccactgcgt	gaatgatatg	20
<210>	43	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of ATEAT1 SSLP market	r ir
	Arabidopsis thaliana subspecies	,
<400>	43	
cgaacagcca	acattaattc cc	22
<210>	44	
<211>	18	
<212>	DNA	
<213>	Artificial sequence	
	• • • • • • • • • • • • • • • • • • • •	
<220>		
<223>	Forward primer for PCR amplification of NGA63 SSLP marker	in
	Arabidopsis thaliana subspecies	
<400>	44	
aaccaaggca	cagaagcg	18
<210>	45	
<211>	18	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Payerse primer for DCD amplification of MCLC2 core	
~6637	Reverse primer for PCR amplification of NGA63 SSLP marker Arabidopsis thaliana subspecies	ın

<400>	45	
acccaagtga t	cgccacc	. 18
		,
		3
<210>	46	
<211>	21	
<212>	DNA	
<213>	Artificial sequence	
<220>		•
<223>	Forward primer for PCR amplification of NGA248	SSLP marker in
	Arabidopsis thaliana subspecies	
<400>	46	
taccgaacca a	aaacacaaag g	21
<210>	47	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA248	SSLP marker in
	Arabidopsis thaliana subspecies	
400	47	
<400>	47	
reteratore d	ggtgaattet ee	22
tergeatere s	ggegaactee ee	22
<210>	48	
<211>	22	
<212>	DNA ·	
<213>	Artificial sequence	
(213)	ALLILIUI Sequence	
<220>	·	
<223>	Forward primer for PCR amplification of NGA128	R SSLP marker in
(223)	Arabidopsis thaliana subspecies	, sper marker an
	Alabidopsis thailana subspecies	
<400>	48	•
(400)	10	
gatetattaa	tgtcgtaagt cg	22
agecegeeya '	-333	
<210>	49	
<211>	22	
<211>	DNA	
		,
<213>	Artificial sequence	
<220>		
4440 2		

WO 99/19492 PCT/EP98/06977

<223>	Reverse primer for PCR amplification of NGA128 SSLP marker in Arabidopsis thaliana subspecies
<400>	49
atcttgaaac c	Etttagggag gg 22
<210>	50
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA280 SSLP marker in Arabidopsis thaliana subspecies
<400>	50
ctgatctcac g	gacaatagt gc 22
<210>	51
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA280 SSLP marker in Arabidopsis thaliana subspecies
<400>	51
ggctccataa a	aagtgcacc 20
<210>	52
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	·
<223>	Forward primer for PCR amplification of NGA111 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	52
ctccagttgg a	agctaaagg g 21
<210>	53
<211>	21
<212>	DNA
<213>	Artificial sequence

<220>	
	Povovo primor for DCDlification of VCD111 CCID well-
<223>	Reverse primer for PCR amplification of NGA111 SSLP marker in
	Arabidopsis thaliana subspecies
.400-	
< 400>	53
	1202221 0 00
tgttttttag g	Jacaaatggc g 21
<210>	54
<211>	20
<211>	DNA
<213>	•
(213)	Artificial sequence
<220>	
<223>	Demind primer for DCD analytication of MCDate Control
(223)	Forward primer for PCR amplification of NGA168 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	54
(400))
CCFFC3C3FC C	222.2266636
ccttcacatc c	caaaacccac 20
<210>	55
<211>	20
<212>	DNA
<213>	Artificial sequence
(213)	Altilitial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA168 SSLP marker in
	Arabidopsis thaliana subspecies
	ablaopolo diallana babbpeeteb
<400>	55
gcacataccc a	caaccagaa 20
J = = = = = = = = = = = = = = = = = = =	
	•
<210>	56
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA1126 SSLP marker
-	in Arabidopsis thaliana subspecies
<400>	56
cgctacgctt t	tcggtaaag 20
J	29

<210>	E7 .
	57
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA1126 SSLP marker
	in Arabidopsis thaliana subspecies
<400>	57
gcacagtcca	agtcacaacc 20
<210>	58
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA361 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	58
aaagagatga	gaatttggac 20
<210>	59
<211>	23
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA361 SSLP marker in Arabidopsis thaliana subspecies
<400>	59
acatatcaat a	atattaaagt agc 23
<210>	60
<211>	18
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA168 SSLP marker in Arabidopsis thaliana subspecies
.400-	

tcgtctactg	cactgccg
<210>	61
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA168 SSLP marker in Arabidopsis thaliana subspecies
<400>	61
gaggacatgt	ataggagcct cg 22
<210>	62
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of AthBIO2 SSLP marker in Arabidopsis thaliana subspecies
<400>	62
tgacctcctc	ttccatggag 20
<210>	63
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	Description of Athron CCLD marker
<223>	Reverse primer for PCR amplification of AthBIO2 SSLP marker in Arabidopsis thaliana subspecies
<400>	63
ttaacagaaa	cccaaagctt tc 22
<210>	64
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	Parand animar for DOD amplification of achimious coin marks
<223>	Forward primer for PCR amplification of AthUBIQUE SSLP marke in Arabidopsis thaliana subspecies

<400>	64
aggcaaatgt	ccatttcatt g
<210>	65
<211>	20
<212>	DNA
<213>	Artificial sequence
;<2 2 0>	
<223>	Reverse primer for PCR amplification of AthUBIQUE SSLP marker in Arabidopsis thaliana subspecies
<400>	65
acgacatggc	agatttctcc 20
<210>	66
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA172 SSLP marker in Arabidopsis thaliana subspecies
<400>	66
agctgcttcc	ttatagcgtc c 21
<210>	67
<211>	19
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA172 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	67
catccgaatg	ccattgttc 19
<210>	68
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	

<223>	Forward primer for PCR amplification of NGA126 SSLP marker in Arabidopsis thaliana subspecies
<400>	53
gaaaaaacgc t	actttcgtg g 21
	•
<210>	6 9
<211>	22
<212>	DNA
<213>	Artificial sequence
•	
<220>	
<223>	Reverse primer for PCR amplification of NGA126 SSLP marker in Arabidopsis thaliana subspecies
<400>	69
caagagcaat a	tcaagagca gc 22
<210>	70
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA162 SSLP marker in Arabidopsis thaliana subspecies
<400>	70
catgcaattt g	catctgagg 20
<210>	71
<211>	22 .
<212>	DNA
<213>	Artificial sequence
<220>	·
<223>	Reverse primer for PCR amplification of NGA162 SSLP marker in Arabidopsis thaliana subspecies
<400>	71
ctctgtcact c	ttttcctct gg 22
•	
<210>	72
<211>	21
<212>	DNA
<213>	Artificial sequence

<220>				•		
<223>	Forward primer for PCR amplification Arabidopsis thaliana subspecies	of	NGA6	SSLP	marker	in
<400>	72					
tggatttctt	cctctcttca c					21
-210						
<210>	73					
<211>	21					
<212>	DNA					
<213>	Artificial sequence					
<220>						
<223>	Reverse primer for PCR amplification Arabidopsis thaliana subspecies	of	NGA6	SSLP	marker	in
<400>	73					
atggagaagc	ttacactgat ć					21
<210>	74					
<211>	20					
<212>	DNA					
<213>	Artificial sequence					
(223)	Altificial sequence					
<220>						
<223>	Forward primer for PCR amplification Arabidopsis thaliana subspecies	of	NGA12	SSL	P markeı	rin
<400>	74 .					
aatgttgtcc	tcccetcctc					20
210						
<210>	75					
<211>	22					
<212>	DNA					
<213>	Artificial sequence					
<220>					•	
<223>	Reverse primer for PCR amplification Arabidopsis thaliana subspecies	of	NGA12	SSL	P marker	in
<400>	75					
tgatgctctc	tgaaacaaga gc					22

<210>	76
<211>	21
<212>	DNA
<213>	Artificial sequence
	·
<220>	·
<223>	Forward primer for PCR amplification of NGA8 SSLP marker in
	Arabidopsis thaliana subspecies
	Manager Charlana Subspecies
<400>	76
gagggcaaat	ctttatttcg g
5 555	2
<210>	77
<211>	22
<212>	DNA
<213>	Artificial sequence
	ozracia i bogaciico
<220>	
<223>	Reverse primer for PCR amplification of NGA8 SSLP marker in
	Arabidopsis thaliana subspecies
	asiaopsis enaliana subspecies
<400>	77
tagettteat	ttataaacat cc 2
- 55 5 -	2.
<210>	78
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA1107 SSLP marker
	in Arabidopsis thaliana subspecies
<400>	7,8
gcgaaaaaac	aaaaaaatcc a
	·
<210>	79
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA1107 SSLP marker
	in Arabidopsis thaliana subspecies
-400>	7.0

cgacgaatcg	acagaattag g	-		21
<210>	80			
<211>	21			
<212>	DNA			
<213>	Artificial sequence			
<220>				
<223>	Forward primer for PCR amplification of NGA225 Arabidopsis thaliana subspecies	SSLP	marker	in
<400>	80			
gaaatccaaa	tcccagagag g	•		21
<210>	81			
<211>	22			
<212>	DNA			
<213>	Artificial sequence		-	
<220>	·			
<223'>	Reverse primer for PCR amplification of NGA225 Arabidopsis thaliana subspecies	SSLP	marker	in
<400>	81			
tctccccact	agttttgtgt cc		:	22
220				
<210>	82			
<211> <212>	19			
<212>	DNA			
(213)	Artificial sequence			
<220>	•			
<223>	Forward primer for PCR amplification of NGA249 Arabidopsis thaliana subspecies	SSLP	marker	in
<400>	82			
taccgtcaat	ttcatcgcc		1	L9
<210>	83			
<211>	22			
<212>	DNA			
<213>	Artificial sequence			
<220>				
<223>	Reverse primer for PCR amplification of NGA249 Arabidopsis thaliana subspecies	SSLP	marker	in

<400>	83
ggatosotaa o	tgtaaaatc cc 22
<210>	84
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	·
<223>	Forward primer for PCR amplification of CA72 SSLP marker in Arabidopsis thaliana subspecies
<400>	84
aatcccagta a	ccaaacaca ca 22
<210>	85
<211>	20
<212>	DNA
<213>	Artificial sequence
	•
<220>	
<223>	Reverse primer for PCR amplification of CA72 SSLP marker in Arabidopsis thaliana subspecies
<400>	85
cccagtctaa c	cacgáccac 20
<210>	86
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA151 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	86
gttttgggaa g	ttttgctgg 20
<210>	
<210> <211>	87 24
<211> <212>	DNA
<213>	Artificial sequence
	ne erreen deduction
<220>	

47

<223>	Reverse primer for PCR amplification of NGA151 SSLP marker i Arabidopsis thaliana subspecies
<400>	87
cagtctaaaa	gcgagagtat gatg
<210>	88
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA106 SSLP marker is Arabidopsis thaliana subspecies
<400>	88
gttatggagt	ttctagggca cg 22
<210>	8 9
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA106 SSLP marker in Arabidopsis thaliana subspecies
<400>	89
tgccccattt	tgttcttctc 20
<210>	90
<211>	
<212>	DNA
<213>	Artificial sequence
<220>	·
<223>	Forward primer for PCR amplification of NGA139 SSLP marker in Arabidopsis thaliana subspecies
<400>	90
agagctacca	gatccgatgg 20
<210>	91
<211>	21
<212>	DNA
<213>	Artificial seguence

<220>		
<223>	Reverse primer for PCR amplification of NGA139 SSLP marker Arabidopsis thaliana subspecies	ir
<400>	91	
ggtttcgttt	cactatecag g 2	1
-230-		
<210>	92	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of NGA76 SSLP marker i Arabidopsis thaliana subspecies	n
<400>	92	
ggagaaaatg	tcactctcca cc 2	2
<210>	93	
<211>	20	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA76 SSLP marker i Arabidopsis thaliana subspecies	n
<400>	93	
aggcatggga	gacatttacg	0
<210>	94	
<211>	20	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of ATHSO191 SSLP marke in Arabidopsis thaliana subspecies	r
<400>	94	
ctccaccaat	catgcaaatg	0

<210>	95	
<211>	21	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of ATHSO191 SSLP mark	er
	in Arabidopsis thaliana subspecies	
<400>	95	
tgatgttgat g	ggagatggtc a	21
_		
<210>	96	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of NGA129 SSLP marker	in
	Arabidopsis thaliana subspecies	
.400		
<400>	96	
t caggaggaa	taaagtgag gg	
ccaggaggaa c	caaagcgag gg	22
<210>	97	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
	•	
<220>		
<223>	Reverse primer for PCR amplification of NGA129 SSLP marker	in
	Arabidopsis thaliana subspecies	
	•	
<400>	97	
cacactgaag a	tggtcttga gg	22
<210>	98	
<211>	8062	
<212>	DNA	
<213>	Arabidopsis thaliana ecotype Columbia	
<220>		
<223>	Genomic DNA sequence of AtMSH6	
400		
<400>	97	
ceeeeggee g	ctaacaata aaggtatacg gttttatgtc atcaatataa ctatatataa	60

aagaaatgaa	agatatatat	tgtttttca	tttatcaaac	aaaacaacaa	gactttttt	120
ttacttttta	cattggtcaa	caaaatacaa	gataaacgac	atcgtttaat	catttcccaa	180
ttttacccct	aagtttaaca	cctagaacct	tctccatctt	cgcaagcaca	gcctgattag	240
gaacagcttt	accattctca	tattcctgaa	ctacctgagt	cctctcattg	atctgtttcg	300
ccaaatccgc	ttgtgacatc	ttcttctcca	atctcgcttt	ctgtatcatc	aacctcacct	360
ctgctttcac	acgatccatc	gccgcaggct	ctgtttcttc	ttccagcttc	ttcgtgttaa	420
tcaccggaac	cgccgtagat	ttcccctttt	tgttcgaacc	ggcatcgaat	ttcttaaccg	480
tttgaaccgc	gacaccgttt	ctcagagctg	cgttaaccgc	tttcggatcg	cgtaggtctt	540
ggctcttttg	ttttgatttg	tggagaacta	ctggttccca	gtcttgtgtt	actgctcctg	600
ggtatctgct	cggcatcgtc	gatgaattga	gagaaaggaa	caacgcgaaa	attttattaa	660
tctgagtttt	gaaattgaga	aacgatgaag	atgaagaatg	ttgttgagag	gattgtgata	720
tttatatata	cgaagattgg	tttctggaga	attcgatcat	ctttttctcc	attttcgtct	780
ctggaacgtt	cttagagatg	attgacgacg	tgtcattatc	tgatttgcag	ttaaccaatg	840
ctttttgggt	tggattcgtg	gtacaccata	ttatccgatt	tggctcaatg	gttttatata	900
aatttggttt	tcggttcggt	tatgagttat	cattaaaatt	aagctaacca	aaaattttcg	960
taaaatttat	ttcggtttca	attcggatcc	cttacttcca	gaaccgaatt	attcgaaacc	1020
ggggttagcc	gaaccgaata	ccaatgeetg	attgactcgt	tggctagaaa	gatccaacgg	1080
tatacaataa	tagaacataa	atcggacggt	catcaaagcc	tcaaagagtg	aacagtcaac	1140
aaaaaagtt	gagccctgag	gagtatcgtt	tccgccattt	ctacgacgca	aggcgaaaat	1200
ttttggcgcc	aatctttccc	ccctttcgaa	ttctctcagc	tcaaaacatc	gtttctctct	1260
cactctctct	cacaattcca	aaaaatgcag	cgccagagat	cgattttgtc	tttcttccaa	1320
aaacccacgg	cggcgactac	gaagggtttg	gtttccggcg	atgctgctag	cggcgggggc	1380
ggcagcggag	accacgattt	aatgtgaagg	aaggggatgc	taaaggcgac	gcttctgtac	1440
gttttgctgt	ttcgaaatct	gtcgatgagg	ttagaggaac	ggatactcca	ccggagaagg	1500
ttccgcgtcg	tgtcctgccg	tctggattta	agccggctga	atccgccggt	gatgcttcgt	1560
ccctgttctc	caatattatg	cataagtttg	taaaagtcga	tgatcgagat	tgttctggag	1620
agaggtacta	atcttcgatt	ctcttaattt	tqttatctt	agetggaaga	agaagattcg	1680

tgtaatttgt	tgtattcgtt	ggagagattc	tgattactgc	attggatcgt	tgtttačaaa	1740
ttttcaggag	ccgagaagat	gttgttccgc	tgaatgattc	atctctatgt	atgaaggcta	1800
atgatgttat	tcctcaattt	cgttccaata	atggtaaaac	tcaagaaaga	aaccatgctt	1860
ttagtttcag	tgggagagct	gaacttagat	cagtagaaga	tataggagta	gatggcgatg	1920
ttcctggtcc	agaaacacca	gggatgcgtc	cacgtgcttc	tegettgaag	cgagttctgg	1980
aggatgaaat	gacttttaag	gaggataagg	ttcctgtatt	ggactctaac	aaaaggctga	2040
aaatgctcca	ggatccggtt	tgtggagaga	agaaagaagt	aaacgaagga	accaaatttg	2100
aatggcttga	gtcttctcga	atcagggatg	ccaatagaag	acgtcctgat	gatccccttt	2160
acgatagaaa	gaccttacac	ataccacctg	atgttttcaa	gaaaatgtct	gcatcacaaa	2220
agcaatattg	gagtgttaag	agtgaatata	tggacattgt	gcttttcttt	aaagtggtta	2280
gtaactatta	atctagtgtt	caatccattt	cctcaatgtg	atttgttcac	ttacatctgt	2340
ttacgttatg	ctcttctcag	gggaaatttt	atgagctgta	tgagctagat	gcggaattag	2400
gtcacaagga	gcttgactgg	aagatgacca	tgagtggtgt	gggaaaatgc	agacaggtaa	2460
attagttgaa	acaactggcc	tgcttgaatt	attgtgtcta	taaattttga	caccaccttt	25 2 0
tgtttcaggt	tggtatctct	gaaagtggga	tagatgaggc	agtgcaaaag	ctattagctc	2580
grgggraagg	gaaccatcat	actttatgga	attcgtttac	tgctacttcg	gctaggattt	2640
aagaaatgga	aatcacttca	agcatcatta	gttaggatcc	tgagaactca	ggatgttttc	2700
ttattcgtta	tataataagt	cttttcatca	aggagtaaca	aacaaaactt	gcacaatatt	2760
tgtgtgctca	ctggcaaggc	atatataccc	agctaacctt	tgctagttca	ctgtagtaac	2820
agttacggat	aatatatgtt	tacttgtatg	tggtaccctc	attttgtctc	tcatggaggc	2880
tttcaagcct	tgtgttgaaa	ctggatagtt	acatatgctt	ccaacagaaa	ctagcatgca	2940
gattcatatg	ctttcctatt	ctactaatta	tgtattgaca	cactcgttgt	ttcttttgaa	3000
agatataaag	ttggacgaat	cgagcagcta	gaaacatctg	accaagcaaa	agccagaggt	3060
gctaatactg	taagttttct	tggataggtc	aaggagagtg	ttgcagactg	tttttgatca	3120
tttcttttc	tgtacattac	tttcatgctg	taattaactc	aatggctatt	ctggtctgat	3180
tatcagataa	ttccaaggaa	gctagttcag	gtattaactc	catcaacagc	aagcgaggga	3240
aacatcgggc	ctgatgccgt	ccatcttctt	gctataaaag	aggtttgtta	tttacttatt	3300

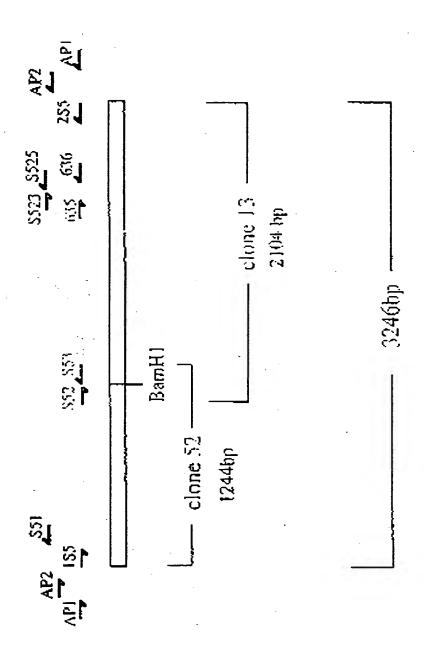
tatcttatca	tgttcagttc	atccaagtcc	tgaaaaatta	cactcttctt	taccaatctt	3360
ccatcaagct	gtgtaaagga	tttggaatta	gaaaatcatt	atttgatgct	ttgttttata	3420
tgcaagaggt	tcccttgaaa	agatctgttt	aagattcttt	gcacttgaaa	aattcaatct	3480
ttttaagtga	atcccctact	ttcttacaat	gatcatagtc	tgcaattgca	tgtcaagtaa	3540
tatcattcct	tgttactgca	tccccctctt	tcttaatgac	cattgtctat	gttgtgtttg	3600
tctcgtgtgc	tggagaaaat	gatagctgat	ccaagctgta	cattatcatg	attaagtagc	3660
tgctcaggaa	ttgcctttgg	ttacattgcc	taatggtttg	atgtcaattt	ttcttctgaa	3720
tctttattt	agatcaaaat	ggagctacaa	aagtgttcaa	ctgtgtatgg	atttgctttt	3780
gttgactgtg	ctgccttgag	gttttgggtt	gggtccatca	gcgatgatgc	atcatgtgct	3840
gctcttggag	cgttattgat	gcaggtaagc	aagtgtattc	tgtatcttat	gtgtaccatg	3900
tgacttcctg	tgcatatatt	tgggttgcag	gaactaattc	tgaatcacca	tttggtatgt	3960
tttttccagg	tttctccaaa	ggaagtgtta	tatgacagta	aaggtaaact	gcttgtatcg	4020
ccagttgttt	tgttaaacag	aatttaaggt	aaatgacact	ggttaattta	aagtgcatac	4080
atgttgaaat	attgcagggc	tatcaagaga	agcacaaaag	gctctaagga	aatatacgtt	4140
gacaggtacc	atttcagtag	gcaagctaac	tgacaattta	accgctcacc	gaatgatagg	4200
tctcttaaac	attgctaatg	tagatgatgt	ttatgtttca	atctaatagg	gtctacggcg	4260
gtacagttgg	ctccagtacc	acaagtaatg	ggggatacag	atgctgctgg	agttagaaat	4320
ataatagaat	ctaacggata	ctttaaaggt	tcttctgaat	catggaactg	tgctgttgat	4380
ggtctaaatg	aatgtgatgt	tgcccttagt	gctcttggag	agctaattaa	tcatctgtct	4440
aggctaaagg	tgtgttggct	tgtttagttt	ttgcttttca	caaattaagc	aaaggaactt	4500
ttcataactt	acagtttcta	tctacttgca	gctagaagat	gtacttaagc	atggggatat	4560
ttttccatac	caagtttaca	ggggttgtct	cagaattgat	ggccagacga	tggtaaatct	4620
tgagatattt	aacaatagct	gtgatggtgg	tccttcaggc	aagtgcatat	ttcttttttg	4680
ataacttcaa	ctagagggca	gacatagaag	gaaaaattct	aatacttcgt	acggatctcc	4740
agtaagtaat	agccgatttt	tgtttaccta	tgtagggacc	ttgtacaaat	atcttgataa	4800
ctgtgttagt	ccaactggta	agcgactctt	aaggaattgg	atctgccatc	cactcaaaga	4860
tgtagaaagc	atcaataaac	ggcttgatgt	agttgaagaa	ttcacggcaa	actcagaaag	4920

catgeaaate	actygecage	atctccacaa	acttccagac	ttagaaagac	tgctcggacg	4980
catcaagtct	agcgttcgat	catcageete	tgtgttgcct	gctcttctgg	ggaaaaaagt	5040
gctgaaacaa	cgagtaagta	tcaatcacaa	gttttctgag	taatgccttc	catgagtagt	5100
ataggactaa	aacattacgg	gtctagctaa	agactgttct	ccttcttttg	caatgtctgg	5160
ttattcatta	catttctctt	aacttattgc	attgcaggtt	aaagcatttg	ggcaaattgt	5220
gaaagggttc	agaagtggaa	ttgatctgtt	gttggctcta	cagaaggaat	caaatatgat	5280
gagtttgctt	tataaactct	gtaaacttcc	tatattagta	ggaaaaagcg	ggctagagtt	5340
atttctttct	caattcgaag	cagccataga	tagcgacttt	ccaaattatc	aggtgcccat	5400
ctatctttca	tactttacaa	caaaatgtct	gtcactactc	aaagcaatgc	atatggctta	5460
gatctcaact	cacaccccga	ggatcctaaa	gggatttgct	ttttattcct	aatgtttttg	5520
gatggtttga	tttatttcta	acttgaactt	attaatcttg	taccagaacc	aagatgtgac	5580
agatgaaaac	gctgaaactc	tcacaatact	tatcgaactt	tttatcgaaa	gagcaactca	5640
atggtctgag	gtcattcaca	ccataagctg	cctagatgtc	ctgagatctt	ttgcaatcgc	5700
agcaagtctc	tctgctggaa	gcatggccag	gcctgttatt	tttcccgaat	cagaagctac	5760
agatcagaat	cagaaaacaa	aagggccaat	acttaaaatc	caaggactat	ggcatccatt	5820
tgcagttgca	gccgatggtc	aattgcctgt	tccgaatgat	atactccttg	gcgaggctag	5880
aagaagcagt	ggcagcattc	atcctcggtc	attgttactg	acgggaccaa	acatgggcgg	5940
aaaatcaact	cttcttcgtg	caacatgtct	ggccgttatc	tttgcccaag	tttgtatact	6000
gttagataa	ttactctatt	ctttgcaatc	agttcttcaa	catgaataat	aaattctgtt	6060
tctgtctgc	agcttggctg	ctacgtgccg	tgtgagtctt	gcgaaatctc	cctcgtggat	6120
actatcttca	caaggcttgg	cgcatctgat	agaatcatga	caggagagag	taagttttgt	6180
ctcaaaata	ccaattcctc	gaactattta	ctcagatttt	gtctgattgg	acaaggtggt	6240
ttgctttt	tttaggtacc	tttttggtag	aatgcactga	gacagcgtca	gttcttcaga	6300
atgcaactca	ggattcacta	gtaatccttg	acgaactggg	cagaggaact	agtactttcg	6360
itggatacgc	cattgcatac	tcggtaacct	gctcttctcc	ttcaacttat	acttgttgat	6420
caacaaaac	atgcaattca	ttttgctgaa	acttattgat	ttatatcagg	tttttcgtca	6480
ctqqtagag	aaaqttcaar	atcagatact	CEEEggaage	cattaccacc	CECECACCAA	6540

ggaattcgcg	tctcacccac	gtgtcacctc	gaaacacatg	gcttgcgcat	tcaaatcaag	6600
acctgattat	caaccacgtg	gttgtgatca	agacctagtg	ttcttgtacc	gtttaaccga.	6660
gggagcttgt	cctgagagct	acggacttca	agtggcactc	atggctggaa	taccaaacca	6720
agtggttgaa	acagcatcag	gtgctgctca	agccatgaag	agatcaattg	gggaaaactt	6780
caagtcaagt	gagctaagat	ctgagttctc	aagtctgcat	gaagactggc	tcaagtcatt	6840
ggtgggtatt	tctcgagtcg	cccacaacaa	tgccccatt	ggcgaagatg	actacgacac	6900
tttgttttgc	ttatggcatg	agatcaaatc	ctcttactgt	gttcccaaat	aaatggctat	6960
gacataacac	tatctgaagc	tcgttaagtc	ttttgcttct	ctgatgttta	ttcctcttaa	7020
aaaatgctta	tatatcaaaa	aattgtttcc	tcgattataa	caagattata	tatgtatctg	7080
tcggtttagc	tatggtatat	aatatatgta	tgttcatgag	attggtcaag	agaaatactc	7140
acaaacagta	tattaagaag	gaaatatgtt	tatgcattaa	tttaagtttc	aagataaact	7200
gcaaataacc	tcgactaaag	ttgcaaagac	caaacacaaa	ttacaaaact	tataagactt	7260
aagttctgaa	ttccctaaaa	ccaaaaaaaa	aaacagaaca	tattttgttg	catctacaaa	7320
caacacaaac	ctacatagtt	tataacttac	tcatcactga	gattaacatc	agaatcattc	7380
tccatttctt	catcttcact	ctcatcatca	tcaccaccac	catgatgatt	ctcctcct	7440
tcacgtaacc	tagcaatctc	actctgagct	ctatcaacaa	tetgettett	ctgcaactcc	7500
aaatctctct	gaaaatcagc	téteatette	tccaactcct	tcatttgctc	tttcttactc	7560
ttctccatct	tctcataaac	cttcccaaac	ctctcaacag	aatccgccaa	catcttatac	7620
gaagcagcgt	cattaacctt	cttcctctcg	tactcaacct	catcatcctc	atcctcctcc	7680
tcttcagaat	caccaggact	atccatcatc	tcatcaaacc	cattagactt	atctaaataa	7740
accttagtgt	tcataaacac	aaactcacct	gaatcaacac	cacaagctaa	acctaaatcc	7800
gacttgggcg	aaacacaaag	caacatatcc	aacttattga	aaaacgacca	tttacttgaa	7860
cctaaacctg	atttctcaac	cttaatcttc	tcttttctat	acttcctctt	caagtcatca	7920
atcattctcc	tacattgcgt	ctcagatttc	tccatcctta	gctcctcact	cactttctca	7980
gctacttcat	tccaatcctc	gttcctcaaa	ctccttctac	ccaattgcaa	aaacctatct	8040
ccccaaactt	caagcaacac	aa				8062

THIS PAGE BLANK (USPTO)

Figure 1



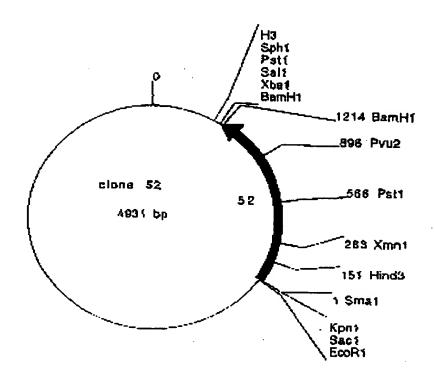


Figure 2

Comments/References: 52= 3' side of S5 (AIMSH3) 1244bp in pUG18/Sma1

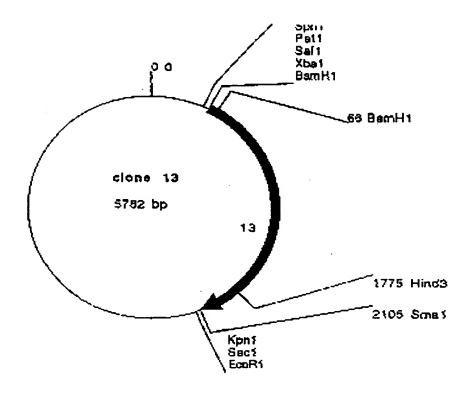


Figure 3

Commenta/References: 13 = 3' side of \$5 (AtMSH3) 2104bp in pUC18/Sme1

Figure	45
Figur	a
Figu	H
Fic	2
1	10.
	Ė

30	144	204 35	264 55	324 75	384 95	444 115	504 135	564 155	624 175	681 195	744 215	804 235	864 255	924 275
MGGA	000 4	900 &	CIC	GAT	GFT V	AAG K	66A 6	TIC	វ្ល	ಕ್ಷನ್ನು ಕ	455 E	aat N	54.89 13.00 10.00 10.0	GTT V
actaagaaracgogogaaattggcaacccaagttogocatagocacgaccacgaccaccacttccattrotottaaacggagga	तिया त	S CO	CAC A	22	TAT (CTA 7	TTC (ALT T	AAT C	ser c	CTF C	ACT A	ATT G	Grt g y
Tar	TIC	000 A	GAC (GTA (सर्य	646 7	TTC 3	CAC	ore 1	Car H	30G (CHG 3	GGT 7	GRA (
PCIC	TTC	ACA T	TCC	CC S	GAG R	CIG V	AGA R		CTG	TCC S	ည္တန	E 5	101 C	GGT .
CATT	CGT B	TCA S	CTC	AHC N) () ()	5 v	TAC	ATG M	₹	A K	A M	G31	<u>ي</u> وو	5
CZIG	TCT	ద్ద క	ti.	4 00	TCG	25. 20.	954	64 (C) 32	15 2 %	AIT I	T T	T T T E	tit 2	s S
ACGA(RAT	क्षेत्र इ	y S	act	သ ရ	0.40 0.00	TS:X	GCT A	ctc v	ς γ.	TRT ï	i S S	ACA T	FIT. I
ರಾಜಾ	acc T	1) 1) 4	CGT R	ပ္ပန္ ပည္သ	ភូមិ ក	433 3	500	TAC Y	n; 1; 1;	ું જુ	TIG	ង្គី	6.4G	GRA S
ង្គ	9 5 0	GT's V	A A B			TTG L	Grr	417 I	7	\$CT	ಎ೧೦ ೩	र जू	TCS S	GTT
atag	30	550	100 S	TCT	TIT	អ្នក	r F P	က လို	ደይች ች	a S	rcg \$	s GCT	argo X	ပ္ပ ပ
ರಿದಿದ	A X	ra? N	r P	ch L	ACA R		ं इ	TTG L		ACT T	cre i.		GT 7	GTT V
AGTT	2 0 0	900	វិជី	X S	ភ្ជួ		A D	GTG V	C 보 라	Š o	5 6	16t C	AG4 R	crt v
4 000	SA X	440 13	TIC	CC-1	ភ្ជ	RAN X	ITG	ည္တမွ	t'I' E	ନ୍ଦ୍ର K	ପ୍ରପ୍ର ନ	ច ច	ន ខ្មែ	69
SCA.	ပ္သည္ မ	SAC R	100 8	A.A.G	TTT		GTT	වී අ	₫ L	GTG	:TC	នច្បា ត	GAT D	GTC V
attg	ATG M	ACT T	GTA	A A	A CA	S	GT >	45 46 47 48		67A	rit E	RGI S	GTG ✔	AGA R
GAAA	gattaccaatabaccaatt	9 U B,	ACT.	200	స్ట్ర ఫ	TCA	GAT D	ATC	GTG V		cc1 r	ATA I	GTT V	GTT V
2525	Pa GC	700 %)) 4	7CA 5	to H	aca S	ದ್ದಿ 🏲	ថ្ម ថ្មី ធ	AGT S	ATT	6 0 0	5. 5. 5.	17. C	car o
AAGC	AATA	2 ×	10C	900 A	TTR L	PCG	TAC *	95 4	000 4	A.A.C	ACC T	S S S	GIT	ttt F
AMGR	TACG) 2 2	ATA I	00 ₹	rat N	GA.	A.F.G	2 4 5	ACG T	TEC	9 9 4	SCT A	TTG	AGT S
Ω.	GAT	AAA K	A.A.G	0 V	22 -) 200	\$ S	মুমুত জ	ភ ាក ភ	GGA A	्र भू	909 4	TTC	R TG
~	 	145	36	265 56	325 76	385 96	445	505 136	565 156	625 176	665 196	745 216	805 236	865 256

984 295	1044 315	1104 335	1164 355	1224 375	1284 395	3344 435	1404	1464 455	1524	1584 495	1644 515	1704 535	1764 555	1824 575
rre L	otc v	AGC S	ITA L	ott V	cat E	S S	AAT N	ora V	ata I	TCC S	F	1600 c	, GCT	R R
AGC S	1 TG	11 TC	N N	ដូ ។	1GC	TTG	A A	ACA	TTG	ភ្ជិ	S C	8 15 8	GAA E	A T T
TIG	TTT	161 C	GCT G	F 1	7.t'l 6	act S	GTG ✓	CIT		AG4 8	CCT	ජ ව ය	à.TG ⊠	8 8 8
ATT I	7. X	GAT	ភ្ជ ភូឌ	35°C	7 7	CS PR	GTT V	r T	₽. ₽.	K K	S S	art I	sati V	act s
ភ្នំ 🔻	୍ୟୁ ଅ ଅ	CTG	ನಿ ನಿ	S	(E)	T T T F	មិន ខ្លួ	S S S	GAT O	e CT S	61'À V	GAT D	ប៊ី ២ 4	o o
504 €	7 C T	TCF S	O Al H	A.TG	3	S S	1:3 L	A.T.	135 0	ត្ត ប្រហ	att 1	TC: S	rta I	න්ය වූ
တ္ဆ ဗိ	500	in Li	स् इ.स. १	3 5 5	။ ဦ ၁	in a	9 3 0	9.56 K	ج. ج	N TG		7C.	11¢	K K Y You
TTA L	130	(65)	GP.R.	a A A A)) (4)	9 9 9	₹ 5 &	ENT M	CCI 6	555	AG. 84	왕 전 전	15.55 15.55	ATH 1 Sont
(50) C	TC3 S	तृ नु प्र स	TGT	សូ ប៉ុន្តែ	కే చేర	₹ 30 0	0 10	C&T ≭	ງ ປ	scr *	6. 19.00 19.00	ICI S	1	့အမ ွာ
2014 S	CTT L	gre v	172 C	ĠĊŢ ≯	GI I V	TAC Y	ACT T	ي ت	a Sign	S	TCT S	ат К)) «	CIT
494	ty a	ය දුරි අ	ų Š	GCT	1 1	נידי נ	£. ₹	ar J	GTG V	ett I	GGT G	0CT	रू इ.स	CGG CTI R L Figure
r K	940	GPT V	E I	646 1	ញ រ	35.4 1	ည် (၁)	55 %	9 (~ %	၌ တို့ ရ	स्वंध ध	Ž,	GCI	0 GRG
11C 6	999	() 6, z	GTT V	CTG L	CAT H	ာ ရ , နှ	S S	ပ္ပ)) (S	eff L	116	ACT T	Art I
RAT R	נינ	ភ្នំ	្ត ទ	arg X	ర్ట్ ఎ	erro L	CIC T	₹¢r s	እርና ብ	GTT V	GPT V	GTC ¥	000 R	0 0
GRI	TTG L	ACC T	GAT O	ATG M	ATG M	TTT E	121 1	क्षेत्र व	CII	&¢∓	TTG	TCP.	CAT M	A.A.G K
rat V	CTG L	CCT	GTA V	्र सु	A.F.C N	503 5	A A A	දේශ	r L	C. T. C.	GAG B	TCC	TT\$ F	ဗ္ဗ
11. TTC	្តិ ខ្	45 15 15 15 15 15 15 15 15 15 15 15 15 15	450 4	A X	4 2	TTT	ુ કુ	ಕ್ಟರಿ ೧	الم 10 م	CIT	AGT S	ដូច	ATC I	90 9
្ត ភូមិ ភូមិ	A SCI	SCT.	z z z	GRT D	ATT I	ე ე	્રું •	gat D	700 5) () ()	S	GTG ❖	e E	L L
GA.	F CCA	E CAT	GGT	GAT	ACA T	A A	S S	10% 8	ខ្លួ	SCI	CIC	CTC L	ACA T	TTA L
TAT	ACP.	4 4 4 4	AAT	r Sept	CAT		AGT	ZAT ZAT	TAT Y	S	ည် လူ	TRT Y	ata I	ATT
925 276	985 296	1045 316	1105 336	1165	1225 376	1285 396	1345 416	1405	1465 456	1525	1585 496	1645 516	1705 536	1765

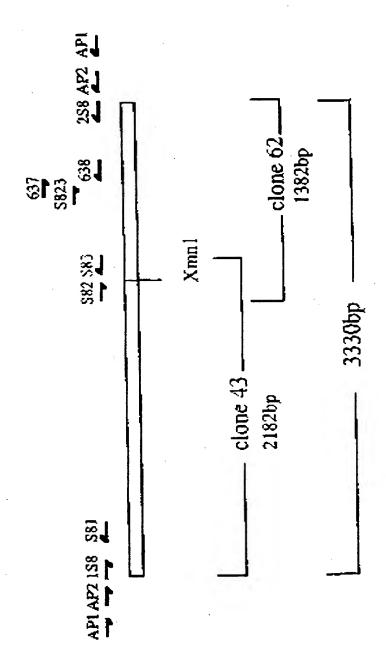
1884 595	1944 615	2004 635	2064 655	2124 675	2384 695	2244 715	2304 735	2364 755	2424 775	2484 795	2544 815	2604 835	2664 855	2724 875
TCB S	GCT	GCT	TTE	TTG	ASG K	g g	AGT S	E Ç	rgr c	cki O	etc I	ATA I	ر د وا	GAA E
ATT	9 ဗ	CIT L	ਰ ਨ	S S S	1. 1.	CTA L	TTC	TTG L	უ მ ი	TTA L	ART	1 1 1	C CAT	r cr
GTT	នុង	GAG D	GCT A	3 ⊢	AGC S	S F	AGT S	161 C	GAT D	ATA I	4	att I	CTT L	1717 F
rcr	ጽ ጽ	CC 4	ATP. I	ATC I	RAT N	CTA	AAG K	G&C D	GTG V	ACT T	755	tir L	GTG V	T T
ATT I	AAT	TTT	7 2 2	ტეტ ტ	45 >	១ ទ	CTC L	ក្នុ ស្ព	TTT E	ည် (၁)	TAT Y	A 601	ည်း ဆ	S.
TTG L	CIR	5 00	3 3 3 3	TCG S	ج و 3	(3)		ස් දුර	GAG E	ភ្ជា	ख रहे ज	GTT V	CTG	AGA R
3 ×	ည်း	ည် က	GAT O	មួ >	613 v	<u>.</u>	agt S	SCT.	ς Γ	হৈ >	396 G	5 0	3.	6 CAT 660 H 6
3 2	عري	194 20	្សា	ର୍ଷ ଜ ଜ	E M		CAT D	CTT.	CGT KGT	ខ្ពុំ	រដ្ឋ ស	r S	209 «	CAT #
TTS	CTC L	TCC	R.S.S. X	ii.	i. 水之	D3 at	¥	हत्त् इ	61.0 V	E E	स ११ व	ATC I	TTČ F	9 0 O
CTI	CTT	ACT T	G A A	TTT	ATG H	ξ;) >	TCG S	ಕ್ಷ ಭ	Tat Y	A F	CET	Tal >	TCA. S	ATC
ACT T	4 ×	atc I	398. R	S F F	111		GCT A	67.T	2.5 ∑	SGT	11G L	750 7	್ರ ಕ	AGT ATC S I
TCT S	800 0	CT3 L	ATC I	TTG	6)C	(2 P) 첫	ig k	353	A Des	TCT	E F	AGC S	<u> </u>	រូវ ប្រ
C63) ()	A:23 1	GTC V	1.5 A 25	े हुन स	۲ ۲۵۲	25 Z	50 P.) A 2	\$3.0 0	ACE.	AAG K	GTP V	TÇ!
ত্য ১	AAT	CAC D	T.Y.	₹ 55 %	7 1)	ည္သ	ctc ^	6830 X	त्र 13 स	ata 1	GAC	SSA G	i Ti	& CT
ACT	ರ್ಷ ೧	ក្ ក្	CTT V	ATT	GAT D	S ×	ATT 1	TT	TCT 5	₹ ≈	AA3 N	<u>క</u> ్తే	TCC \$	9
చే చ	crt v	TTG L	gC3	SCT A	CT: V	TRI Y	000 4	GAT	CTA 1	asa I	ay a	ATG M	ទិន្ន	ATG 8
मुद्रम	GTG	30	¥ 0	2	្ត ប្រ	8 8 8 8	i i	Ž.	ACT.	ට දි වි	GTC V	3	GTT V	88
200		טנ)) (TCA S	GTT V	TTC	84	C. AG	T T
S S	ii U	8 8 8	ដូន	A AG	GAG E	RCT T	GAS I	ξ. Ω	다 다 다	ర్రే 💂	N N	ក ស្វា ស្វា	SCT SP	TTC
AGT S	700 S	CTT.	**************************************)) (ATA	A.A.G	į,	A CH	700 5	G A.a.	GAT	ACC.	ATG M	GIT V
1625 576	1885 596	1945 616	2005 636	2065 656	2125 676	2185 696	2245 716	2305 736	2365 756	2425	2485 796	2545 816	2605 836	2665 856

2784 895	2844 915	2904 935	2964 955	3024 975	3084 995	3144 1015	3204 1035	3264 1055	3324 1075	3397 1082	3458 5	3522 16
TTA L	TTA	ATH	ACA T	GTG V	ر د د د د د د د د د د د د د د د د د د د	h AGA A	GPA E	0 P	ני ני	TGA TITABICTIAACAITATAGCAACTGCAAGGICTTGAICAICTGTAGTIGCG	255 a	
ata I	A CA	5 12 13 13 13 13 13 13 13 13 13 13 13 13 13	CTG	CTT	\$ C2	<u> </u>	13.00 10.00 10.00	cag e	AGA R	TTR(AAT N	AAA
GTT	ភ្ជុំ	100	TAT Y	A A	ATA L	CGT A	S S S S	GAA E	ATC I	TCTG	TCT	IAS AAGESPASEARASAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
CIT	Trat	TAC Y	rcg S	TRT Y	විරිත	GT.9 V	ပ္တပ္သ	rcT s	ag x	ATCB	GAT	ALA.
706 S	ည္သ	CAT H	CTC >	CTA L	9 4	GAG	AGP R	ភ្ជ	၁၀၀	crrs	A TG	AFFE
GET R	r Tr	# CG ₽	Ç ₹	INC Y	S E	ವರ ₹	9 9	e C	SCT A	AGGI	Sys	Sed)
FCT	ម ប្រ	GIC	TAC) ¥ €	is di ox	स् इ. .	5. 5.	Ti.1 S	41T	JGCA	13	Aced Lou
TCT	GT3 t	TTT F	ACS.	GTG ♥	r CC	TTG	Gist.	2. R	\$ % 3	೧೯೯೦	iat tas asas N	CCC GCA TRA AARDSAAASBAAA P A * Figure 4 (Continued)
161 C	661 150	CII	9 9	5 th	GII	X Set X	CAT T	CTG	N TOO	ated	545 5 8	44.4 S
T T	0 1	GTT V	STT	GRI	æ. ¥	j Z	င်္ဗ ပ	ပ္ရမ	9 8 8	CATT	ு ப ஆ	, T&4
AGB R	ည် 🗷	T73	r S	r F	TTT F	S &	ញ្ញុំ ធ <u>ា</u>	500 K	r L	TRA	4 X	က် ရှင်း (၁၈)
ATC	ACA T	TGI	GGT	GAT	S S S	χ Σ	₹00 24	77.ľ	TTA ASG L K	P.B.T.C.	ત.લ જ	0 0 0
ata I	AGC S	ACA R	ಕ್ಷನ್ನು 🙀	T.P.T	TTT	s t	ក្ត ប្រ ល	TTG	TT,	TIL	bia ace aga ala I I R K	۲ ۲
ပ္ 44 ပဲ 🕿	ect T	e ¥	سا سا	> S	3 8 8	at t L	্র ও ও	GAC D			ACT &	S
ည် ရ	ည္ဟ ဗ	47.5 E	కే ప్రత	ပ္သမ္	A R	သည ်း	at G	GGT	ဌ မ ှင အ	fTT É	att a I	ia ia II
ಲ್ಪಿ ಕ	E A	ชา ชา	3. 2. 2. 2. 3. 3. 3. 3. 3. 3. 3. 3. 3. 3. 3. 3. 3.	A A	ည္သန္ န	A E	5 5	د نا	7¹fC ₹	ర్జ్ క	ret a	ATC
GAA E	ជា ក្រុ	r CIS	AGT	. Sa	S S	Ç ₹	ACA T	GCT	గ్రే క	TGT	ATG T	TEC ¥
AGT	CIT	ပို့သ	A TC	X X	CTT L	ata I	ዶል ፣	a S	A A	A CT	H	ijο
TTA	្តិ ស្ត្រី ធ	CAT B	GAG	9 13 0	6 6	7GT C	. 45 ×	ATT	76G 8	బ్రే 🎍	TACTAACT	775
5	GAT	5 5 5 5		TTG	S S S	ACR S	S S S	1 1 1	ر ان ان	X 관	TAC	GTG V
2725 876	785 896	845 916	2905 936	956 956	1025 976	986	3145 1036	3205	3265 1056	3325 1076	3398 3	34.55 55 56

Figure 5

I WENGER	化自己基金 医克里奇 化甲基苯甲基苯甲基苯甲基苯甲基苯甲基苯甲基苯甲基苯甲基苯甲基苯甲基苯甲基苯甲基苯甲	s spense ta spense property as aby and the state of	-D-10457564565645656666666666666666666666666	G. MAGGRESSTOVARAGEN SEVRYERS OF CONTROL FOR VALUE SACHES SERVICE FOR SERVICE	1. Pikaturgo atraocoskurskurko krigoskuv vrno je bisku krigosku krigosku krigosku krigosku krigosku krigosku p 8. mogunice statem krigosku krigosku po 1. ku krigosku cuki prijestija krigosku projestija krigosku projestu p	4 OF OR OR REVERSOR OF THE TANKE TO THE PROPERTY OF THE PROPER	o erspessioner for the transference older drafts - sector the argument best more and the sector best the least of the sector	AND	S COPYCHAIO 2008 MPT HOUNY HENDY JOHN FOR YOU WINDOWN OF YOU WAS AND AND WAS AND WAS AND BURNESS ON SOME OF THE SAME WAS AND	0 2 2 2 2 2 2 2 3 2 4 2 4	is. He ooksykkingverig endydykking pressing fra som an speakrene hennym obereke mora som som beling of de som i It broedskermen som beling in bestande bestande som bestande som over order of the source.	
	to do	2 4 4	24 24 4 4	900		9 A	6 % % % % % % % % % % % % % % % % % % %	# D	- C	0 7 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 g 6.46 4.4	90
	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	- 17 TO 17 T	X 56 7. 56 8	# # # # # # # # # # # # # # # # # # #	K B B D C G C B B B C G C C C C C C C C C C C C			4 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	, , , , , , , , , , , , , , , , , , ,	4 - C B S	大学のの一年では、大学の日本の一年の日の一年の日の一年の日の一年の日本の一年の日本の一年の日本の一年の日本の日本の日本の日本の日本の日本の日本の日本の日本の日本の日本の日本の日本の	74 - 2 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4

Figure 6



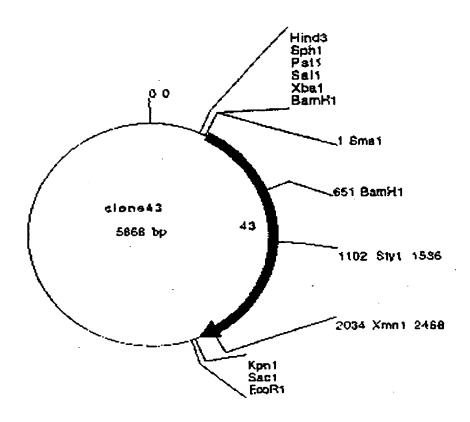


Figure 7

Communits/References: 43. 5 side of S8 (AIMSH6) 2182 by in pUC18/Small

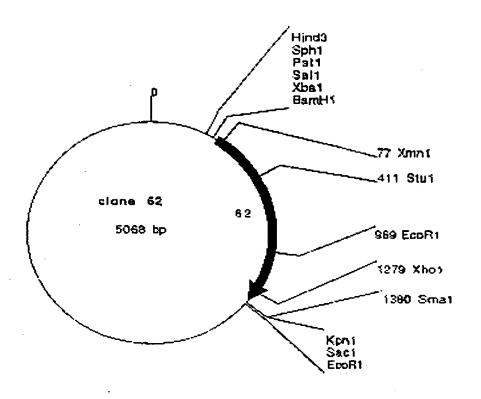


Figure 8

Comments/References: 62= 3' side of S8 (A1MSH6) 1379bp in pUC19/Smet

ð
đ
브
6
-,7

80	153	213	2 4 4 4	333	393	453 104	513 124	573	633 164	693 184	753	813 224	873 244	933
သသ	Q	TCC	ુ આ	AGA R	۵ د درو	₹×	GRT	A A	GAS F	GCT.	GTA	क् यु	AGA R	T T
20	900 000 000	GIT	S. P.	GIT	æg. *	GTA V	rat N	GGT 6	GTA V	CGT R	CCT.	A.A.A.	35. ×	GTT
arcr.	8 8 8	116	A.A.G	9 3 3	TTT	TIT	77.07.0	A LAT	1CA S	8 ~	GT P V	F. K.	8	Gar
3CCF	5 0 0	6 61	516	GAT	1 50 0	A X	ပ္တန္	A Z	aga R	R 061	SE X	GAG E	GAT	6
2002	e E	AAG O	AAT N	GTC	rcr	18. C. 8. T. 18. T. 1	GTT	100 5	CIT	atg 8	GAT D	\$ \$	AGG A	ರ ಕ್ಷ
rtti	TCTCTCAGCTCAAAACAICGITTCTCTCTCTCTCTCTCACAATCCAAAAA ATG	7	TTT	TCT S	ις 13 π	ž Ž	ر: د: ۲	CC5T R	្ត ស	ις υ	049 E	161 C	17	AIA
PARK	TTCC	26 EO E+	СGА 33	44 %	رن د.	ATT	GAT G	1 h	ij) 4	న్ _{గా}	X X	GT.Í	ម ម្រ	တွဲ့ အ
9000	FC.	55	٠ ا	S CG	۲ وا	ልሌፒ ዜ	त्रः इ	4 4 0	AG\$ ∺	şğ L	IT) E	9) a	tcT s	17.E.
GCAA	ICTC)) 	G G	GII	CGT R	70 70 70	§ ≈	rc1 ₽	99 0	ម្ច ម៉ី ស	F.	ពិឝ្តិះ ម	act s	1 1
CGAC	וכוכ	1. 1.	§ 0	100 at	SGT R	TTC	ه 5	att I	AGT S	800 P	atg B	19 20 20	ର ଜୁନ	7. 7.
ICTA	זכאכו	22 ~	S S S	71 71 71 71 71	200 A	CTG L	3.05 R	د د د د	7 7 7	ي در وون	ម្ម ស្វី ធ	υ. 13 13 13 13 13 13 13 13 13 13 13 13 13	CIT	A CA R
CATT	ICIC	ል የ	ပ္တပ္သ ဗ	CGT	GTT	TCC \$	GAG E	GAT U	# (5) (6) (5)	ري ان م	ÇĂ,T O	atg N	TGG 14	CAT.
0000	TTTC	4 6 0	ပ ပ	GT3 V	A A	377. S	& 60 64 64	। भू	T.J.T E	GTT '	រិក្សិ ខ្	44.7 7	ក ស្ត្	TRC
GITE	ATCG	TIC	88 9	TCT	6 46	के दि	FCT	GCT	हुटा इ	GAT D	CIG	ე ექ	ITT E	CTT
TATO	SPAC.	TTC	ည္တ	GCT A	ئ درون م	GAT U	IGI	A X	C&T H	ည ပို့	GTT V	75 7	춫	88
SCAG	CTCA	ror s	ည္ဆင္လ	ပ္မ	<u> </u>	GST G	GAT	A TG	7.8 60	GAT D	5 4	4 4 ×	₽CC T	641 0
CTGA	rcae	2	13 et	y y	i i	ವಿಧಿಧ ಕ	۳ وي ه	า เลา	AGA R	GTA V	A PA	A.A.C.	ទី ៤ ៤	CAT D
DO ST	ıcı	AIT	₽ 6CT	X X	GAT 0	TCC S	GAT	CTA L	स् १५ ७	g g	TTG	s TCI	ig E	ה ק
STIG	Grat.	7CG 8	GAT	SCT SP	7	ស្គី ខ្ម ខ	CAT T	TCT S	a So	ATA I	2 2 2 3 3 4	GAC D	A.F.C	CGT R
aaagtigagoogtgaggastostttcggocalttctacbacbcracgagaaaatttttggggcaatctttcccoo	Titcgrat	AGA A	၁၁၁		క్ష క	GCT	GTC V	ర్డ్ జ	F F			TIG	STP 5	AGA A
-	# T	15.4 5	214	27.24. 25.55.	334 65	394 85	454 105	514 125	574 145	634 165	694 185	754 205	814	874 245

993 28¢	1053 304	1113	1173 344	1233 364	1293 384	1353	1413 424	1473	1533	1593 484	1,653 504	1713 524	2773 544	1633
TT.	667 0	و و و	GIT V	ata A		TCA S	A L	689 14	ACG T	GAT	TCA S	686 B	TII	5 3 3
· `	TTR L	GTT V	¥ ¥	ACT 1	ATC 1	TGT	35°	AAG K	Tat Y	ACA T	e E	.	ATT	CITY 1
ATG G	C. A.	0 0 0	TAT Y	A AT	AAC ≍	rag K	មួ ២	CCA P	æ *	GAT D	S	7	e g	AAT M
IAT A	ე ე ≉	ig k	13 0	GCT	න් ලි පු	3 0	GTT V	Ş u	₽.GG ₽	909 9	ភ្ជិ	ફે ફે	9 9	GTA V
GRA T	GAT	100 C	CCT TO TA	scr G	CAG E	ដែ	₽	GII	CIA	atg M	00.	S S	Cr. B	atg N
ស ស	CTA	र् _ष ्ट्र स	E III	بى بەر بە	3 60	G E	T T T E	i) Q	i F	GTA	র ই	CTT	8.56 ×	8
A DEA X	ig m	2 12 13 13	TTA	υ Ο «	ر د د	11 Z	00 K	≱7G N	2. A 5. A	3 3 3 3	TIT3 F	ည ပြာ တ	CTT	AT GGC CAG AC G Q T (Continued)
A TTS	रहर ४	57G V	स् इ.च	ત્રું રે×	€ .	iz Z	TTG L	111G L	g g	200 P	₩	GTT V	GTP.	GGC G Onti
get G	7. 1.	GCT C	A R R	G C.S.	લ () (- પ્ર	ATC I	S S S	IIA L	ಸ್ಟ () (*	GTA V	ရှိ ဇွေ	GRT D	GAT	ம் 🗁 வ
70 % 20 %	ည် ည	ភិភិព ភ	ಕ್ಷ ಕ್ಷಾ	ន ប៉ុន្	(C) €	ខ្លុំ	GCT	ტებ ₹	65.0 T	^က ရ	8. 18.	7. 101	ពីអភិ ព	ATT I Web
IAT T X	TRT	ATG M	GTG Y	GAC	ACT T	ag x	Tot	i S S	aga K	Ş,	TCT S	§ မ	CIA	AGA ATT R I Figure
0.00 T	T.L.	S S S S S S S S S S S S S S S S S S S	60 4	S S	TIA	41. 1	GAC	دنا	ري د	÷76 1.	ក្នុង ស្ន	aat N	og ×	י נו
eAG o	A.P.B. K	ATG B	3,50 13	A D F	CTA V	£ \$€			ŧ.	20	ata I	ក ម៉ូ	CTA L	TGT
य र्ह्	999	A.R.C.	GAT	GRA E	200	CII		1.00 4	រុះ ទី	GTA V	型 工 和	567 5	AGG R	65T
ည် ဂ		₹66	ATA I	CTA	GTT V	CTT		TGT	a a	909 4	PAT N	Ç.I	rcr s	35 ×
ಸ್ಥ ಕ್ಷ		OFC O	9 9 9	9 5 0	CTA	これに		ADF.	AGT S	acG T	#5# #	ÇTT	CIG •	7 H.C.
ည် က	TTT E	ព្ឋា	AGT S	9 9	A AG	GTC	Ž.	ე გ	ာ (၃)	TCT S	GTT V	S S S S	S F	GII V
	• •		7. 3.		R PCG) (5)	TRT X	GAT U	13.57 14.17	999 9	608 5	TGF	AAT u	440 O
#	TLO	5. X	rcts	₹ 55 &		GAT	616 •	GAT D	TTA	నే 1	13	A.A.C	ATI	TAC Y
AAG AX	GIG CIT IIC V L F	CAC	ATC	6 6 6	ATT) 1	ACT	8 5	ore.	TTG			्राष्ट्र स	CCA P
934 265		1054	1114	1174 345	1234 365	1294 385	1354	1414	1474	1534	1594 485	3654 505	1714	1774 545

FOCIO «WO 0010402A2TI »

1893 584	1953 604	2013 624	2073	2133	2193 684	2253 704	2313 724	2373 744	2433 764	2493 784	2553 804	2613 824	2673 844	2733 864
A A	GAT O	AGT S	ပ္ပ	GTG V	CAT E	A.A.A.	ರಿ ರ	ACT T	CAC H	450 6	5 L	GGT	att I	R
GAT	* *	CAJ.	4 5 5 5	ess K	att I	161 C	A G	हु <u>ज</u>	att 1	GCT	K A	20	AGC S	r C
CTT	CIC	ភ្នំ ភ	CIC	Æ.A	\$ 5 5 5	CTC 1	ž u	scr *	GTC V	TCI	5 5 5	8 4	ဂ ဝ ဂ	27
Tat Y	CCA P	Mac N	CIG L	0 0 0	rgt S	ž×	17C	₹	ម្ភា ម	ည်သ	PAT N	5 ₹	AGT	ACT T
ž X	t T	60.3 8.	ag a	ر. ا ا	र १५ १५	TAT Y	දී ය	1 30 ω	rcr s	A GT 5	D D	GTT V	36C 8	5 4
TAC	i i	ACG I	4 5 u	C:1	₹. ;•	C.T.	TCT is	0.64 0	766 E	გე დ	SKT D	5 4	AGA R	X X
TTG L	a IC	TTC	4T.	(; (; (;	<u>ب</u> ب	115 1	C:T C	¥Ç¥ €	ક ડે ૭	500 es	7 7	Til	ACA R	Services
A CC	7GC 8	स्ट इंट	ر د و و	٠. ن ع.	\$ %	۱. چ	T I I	515 716	a Ct	14 14 17	sct P	<u>د</u> ۲	DO 4	ეეს ეეს ეეს
999	ይ ት ፓ	લું કું ઉ	n a	13 	013 ¥	् स् स	F1	GÀT D		Í 🛪	4	CAT R	GAG X	ATG B G
చై క	766 7	GTT	CIT	516	AT.	0 1: 4: 21	ក្នុង ខ្	200	(§ 4	TT1 F	ភ្ជី ស	7.00 21	ပ္ပ ပ္ပ	CCA AAC P N Figure
17.0	TTA	ers *	X Ag	ior S	9 20 0	# ¥ =	CJA L	کار تاریخ	4 3 3 3	₹ S	ង្គ ខ	CTS L	ಣ್ಣ	
90	CTC	GET D) 12 13 13 13 13 13 13 13 13 13 13 13 13 13	00 K	ig O	្តិ	့် (၁) (၁)	9 50 50	2 E	at of	8 ~	ង ១១ ១);]	9 9
តិ តិ	S S S	CJ.	01.0 1	ရ သ	TTT F	e G	350 S	ፒዱን Y	TTI	CTG L	TTT F		ATB I	ACG T
GAT D	X AG	13 C) 44	TRT X	Ę N	्रं स	X X	ች ት	異る	CT J	GTC V	ATT I	ATC	GAT	ភព
ig E	GGT	X X	2 2 2 3	₹ ~	4. X	:: 3 a	SGA S	رن ه	3 3	GŘT O	GTT V	ž Ž	ZZT H	77.8 1
S SGC	T T	AAT	ပ္သမ္	GT T V	STT V	CT.	STP >	Tiri F	PATC I		i S		ပ္သ	TTĞ L
PAT	5	ATC	मुद्रा	AGC S	ه د د د	SCT A	ITA L	ე წე	CTT L	ក្តិ	256 R	ATA	GIT V	S S
MAC H			ATC	ICT 5	300	73G L	ATA I	25 2	ATA I	ည်း လ) () ()	500 th	CCT P	9 %
TTT	GTT	ANG R	కే చ	A X	አላፋ ጽ	TTG L	CCT P		a a	ATA J	atg K	ମୁଗ୍ରନ ଜ	1 13	E F
ATA	i Ti	618 V				CTG			r L	ACC T	AGC S	ara X	4 50	CAT
3834 565	1894 585	1954 605	2014 625	2074 645	2134 665	2194 685			2374	591 165	2494 785	\$68 \$68	2614 825	2674 845

•				_		~ -	m	L) S	10. 4	L) de	m 9	-	Ø	9	
2793 884	2853 904	2913 9 24	2973 944	3033 964	3093 984 9	3153	3213 1024	3273 1044	3333 1064	3393 1084	3453 1104	3521 5	3579 1.9	3606 28	
i√ 40	ଲ ବ	9	13 D	m bx	ים נים ניי										
ט	4	Ħ	TAC	TTT	ABA K	ည်း ဝ	Gre v	ည ဗ අ	AGT S	ರ ಕ	ក្ត	ζŢ			
7.00 D	4 t	A ACT			i≹ 3≤ ∪	ಚ∆ ಇ		ড ব্			3 S	ក្ន	A A A		
101 8	ATG	ઇ જ	તું જ			නු වී රූ	ල් ට්ට	₹ 3 0	S S	RAT R	e E N	160	Ę.		
CAG E	ATC	rat v	GRT D				CIT	GCT	TTC	AAC N	AMA K	TIT	450 K		
រាធិរុ ព	AGA R	30	TTC	000 8			క్రో ఆ	GCT A	945 E	D E	3 ATC I	PAGT C	្ត្រី រូប		
ರ ೯	GAT O	1	ACT T	उद्धा	2 K	ភូពិ ព	TAC	CGT	17.77 S	229€	0.40 8	ACA TAA CACTATCIGAAGCICGTTAAGICTTTIGCCICTCI T	TTT (E		
GTG V	TCT S	6.17 <	AGT S	ಕ್ಷ. ೧ ರ	8 34	CGT	39.8 S	7. S	न् अ	CTC v	# £3±	3CCT	1733	. 9	•
7.8℃ >	ಕ್ಷನಿಧಿ ಕ	င် (၁	¥0.4	STT >	0 30 20 20 20 20 20 20 20 20 20 20 20 20 20	رن 10 م	ରୁ ଅ	स्तुष्ट स्तु	CJ'A L	#55 K	r rgg	TGP	4.3. x	of (Continued)	i
760 0	ם ט	97.4	499 5		N A	3 00	CCT	#C# 1	ನಿಸಿದ ಕ	TCT	TTR 1	TRE	F. A. A.	Cont	
ეკე ტ	CIT	4 1 1 1	40k R	S S S	SS ≪	Tr. Y	16T	ନ ଜୁଣ୍ଡ ଅ	a S	art I	76C	3	ATC I	•	
110	75 75 75 75	ស្លួ	္မွာ ဗ	GTP >	TTC ₹	GA? D	GCT A	GT:E	2Ct S	667	TTT	TAH	TAT Y	91	7
30	#CB	ACT T	ខ្មែរ	CTG.	g g a	रदा इ	SGA G	(5) (1)	₹ [%]	GTS V		i e	TTA 7	_	1
55 A	TTC P	166.3) 435 2	CAC X	S X	45.53 55.53	GAG E	5.5 5.5 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0	717. F	7. 1.	ACS T	r U	160 T	स् स्	
	atc T I	GAA T		CGT C	F C F	Į,	ACC T) 12.2 2.2 2.2 2.2 2.2 2.2 2.2 3.2 3.2 3.2	AAC N	ម្តី ស្ត្រ	0 0	ect A	ar Ku	R R	
TTT :							TTA A L	- 45,2≥ -{-{		e oka K	TAC (ATG (A.E.	र्युष अ	
ATC	¥ ₹	GTA V		TT	CTC L	4.5 ×				g;≈	ર્સે⊁ છ	ন্ত ক	A A A		
ÇET V	143 0	T. Dr. J.	ATC I	GIT	CCT P	TTC	CGT	ata I			0 0	TAR.	CTT	4. ¥.	
8	GTG A	TT		ا ا ا	Ç&C #	60 8	1AC Y	ក្ ភូមិ	រៈ[ដ	3. F	ag o	ž×	Ħ	A X	
•					TAC (TGC +	TTG.					Ω 2002 M	ర్థు ఆ	4 4 X	
. CTG			, Di⊷ er	# #	E X	F.O	TIC I	ATG G		. A)) (GIT (문교	€	
TGT	5 S	AGT	ដ្ឋ	ស្ល ប្រ ផ	ರಚ್ಚಿಗೆ ಚ	Ç 3 ∢							TTT ?	e:	
ర్మ క	ATC I	GAG.	E Grt D	ratt 1	ACA T	atg K	GTG	. 513 513					្ន	3 2	
GCA 1								1 100 A	ATG	: ភូមិ) D	TAC Y	, O	A.A.	
734 865					3034 965	3094 985	3354	3214	3274	3334	3394	3954	3522	3580 20	

Figure 10

1910 sprencessing and the servence of the contraction of the contracti	ACM 544- P.
50	100-X; 100-X
	10年3年 10年 10年 10年 10年 10年 10年 10年 10年 10年 10
** * * * * * * * * * * * * * * * * * *	
AND ALLES DES DES SERVICES DE SANTENCION DE	1010 1010
14.2 MINITERINATION OF THE PROPERTY OF THE PRO	75 - 72 - 74 - 74 - 74 - 74 - 74 - 74 - 74
448 RVILLOSSAVORAVORAVORASONEN OVERSONEN OVERSONEN OVERSONES OF STANDARD OF STANDARD OF STANDARD OVERSONES OF	PARA TIME
7 F	1
112 PFFERMINGSTROMENT OF COMPANY CONTRACTORS NAMED TO A CONTRACTOR OF THE PARTY OF	
161 VERLOVEGEVERVERVERENGELENGELENGELEURFELENGEN AURELUERKORGERVERVER 1800 - HELDER PORKER BEN HELDER EN VALLER	14. 14. 14. 14. 14. 14. 14. 14. 14. 14.
6) Isemperations and Active Folding prayed postator postator between the proposition of the same states and the same states are same states are same states and the same states are same states are same states are same states and the same states are same same states are same states are same same states are same states are same states are same same same same same states are same same same same same same same sam	
- ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	10年の記録・10日の日本の日本の日本の日本の日本の日本の日本の日本の日本の日本の日本の日本の日本

PETTTGGTTGGTAACAATAAAGGTATACGGTTTTATGTCATCAATATAA	50
TATATATAAAGAAATGAAAGATATATATTTTTTTTTTT	100
PARCARCABACTTTTTTTTACTTTTTACATTGGTCRRCARAATACRA	150
ATARACGACATCGTTTAATCATTTCCCAATTTTACCCCTAAGTTTAACA	200
etagaacutteteettikki ekingeageageagetti	250
CCATTCTCATATTCCTGAACTACCTGAGTCCTCTCATTGATCTGTTTCG	300
CAAATCCGCTTGTGACATCTTCTTCTCCAATCTCGCTTTCTGTATCATC	350
ACCTCACCTCTGCTTTCACACGATCCATCGCCGCAGGCTCTGTTTCTTC	400
TOCAGOTTCTTOSTCTTAATCACCGGAACCGCCGTAGATTCCCCCTTTT	450
TOCAGETTE I IDSTETIATEMOODEN STANDERSTEEL I INDICAGEGACACEGTTT	500
TGTTCGARCCGGCATCGARTTCGGATCGCGTAGGTCTTGGCTCTTTTG	550
TCABAGCTGCGTTAACCGCTTTCGGTTCCCAGTCTTGTGTTACTGCTCCTG	600
GETATCTGCTCGGCATCGTCGATGAATTGAGAGAAAGGAACAACGCGAAA	650
ATTTTATTAATCTCAGTTTTGAAATTGAGAAACGATGAAGATGAAGAATG	700
ATTTTATTAATCTCAGTTTLEAAATTATATACGAAGATTGGTTTCTGGAGA	T50
PTGTTGAGAGGATTETGATATTTCGTCTGGAACGTfCTTAGAGATG	80 0
ATTCCATCATCTTTTCTCCA: ITTCCCCTTOTCCATCTTTTTCGGT ATTGACGACGTGTCATTATCTGATTTGCACTTAACCAATGCTTTTTGGGT	85D
ATTGREGACGTCATTATCTGRITTGCACTTTTGTATATATATATATGGATTCGTTGGTTGGT	900
TGBATTCGTTGTACACCATATTATCACTAAAATTAAGETAACCA AATTTGGTTTTCGTTCGTTATGACTATCACTAAAATTAAGETAACCA	95D
AATTTGGTTTTCGGTTCGGTTTCAATTCGGATCCCGTTACTTCCA AAAATTTTCGTAAAATTTATTCGGTTTCAATTCGGATCCCGTTACTTCCA	1000
AAAATTTTCGTAAAATTTATTTCGGTTTAGCCGAACCGAATACCAATGCCTG GAACCGAATTATTCGAAACCGGGGTTAGCCGAACCGAA	1050
GAACCGAATTATTCGAAACCGGGGT TAACCGTATACAATAATACAACATAA ATTEACTCGTTGGCTAGAAAGATCCAACGCTATACAATAATACAACATAA	1100
ATTENCTCGTTGGUTAGAGAGATCCAACGGTCAACAACAACATT ATCGGACGGTCATCAAAGCCTCAAAGAGTGAACAGTCAACAAAAAAAGTT	1150
ATCGCACGTTATCAAGCCTCACGACGCCAAGCCCAAGCCGAAAAT GACCCTTGAGGACTATCGTCCCCCATTTCTACGACGCCAAGCCGAAAAT	1200
TTTTGGCGCCAATCTTTCCCCCCTTTTCGAATTCTCTCAGCTCAARACATC	1250
TTTTGGCGCCAATCTTTCCCCCCTTTCGAAAAAATGCAGCGCCAGAGAT	3300
GTTTCTCTCTCTCTCTCTCACAAAAAAAAAAAAAAAAAA	1350
CGATTTTGTCTTTCCAARAACCCACGGGGGGGCGRCTACGAAGGGTTTG	1400
GTTTCCGGCGATCCTGCTAGCGGCGGCGGCGGCAGCGGGAGACCACGATTT	1450
AATGTGAAGGAAGGGGATGCTAAAGGCGACGCTTCTGTACGTTTTCCTGT	1500
TTCGAAATCTGTCGATGAGGTTAGAGGAACGGATACTCCACCGGAGAAGG	1550
TTCCGCGTCGTCTGCCGTCTGGATTTAAGCCGGCTGAATCCGCCGGT	1,600
CATECTTCGTCCCTGTTCTCCAATATTATGCATAAGTTTGTAAAAGTCGA	1650
TGATCGAGATTGTTCTGGAGAGAGGGTACTAATCTTCGATTCTCTTAATTT	1700
TGTTATCTTTAGCTGGAAGAAGAAGATTCGTGTAATTTGTTATTCGTT	1750
GBAGAGATTCTGATTACTGCATTGGATCGTTGTTTACAAATTTTCAGGAG	7600
CCGAGAAGAYGTTGTTCCGCTGAATGAYTCATCTCTATGTATGAAGGCTA	1850
ATGATGTTATTCCTCAATTTOGTTCCAATAATGGTAAAACTCAAGAAAGA	1900
AACCATGCTTTTAGTTTCAGTGGGAGAGCTGAACTTAGATCAGTAGAAGA	1950
TATAGGAGTAGATGGCGATGTTCETGGTCCAGAAACAOCAGGGATGCGTC	2000
CACGTGCTTCTCGCTTGRAGCGAGTTCTGGAGGATGAAATGACTTTTAAG	2050
GREEATAAGGTTCCTGYATTGEACTCTAACAAAAGGCTGRAAATGCTCCA	2100
GGATCCGGTTTGTGGAGAGAAGAAGAAGTAAACGAAGGAACCAAATTTG	2150
AATEGETTGAGTCTTCTCBAATCAGGGATGCCAATAGAAGACGTCCTGAT	2200
CASCCCCTTTACGATAGARAGACCTTACACATACCACCTEATGTTTTCAA	

	2250
GARAATGTCTGCATCACAAAAGCAATATTGGACTGTTAAGAGTGAATATA	2300
TGGACATTGTGCTTTTCTTTAAGGTGGTTAGTAACTATTAATCTAGTGTT	2350
CANTOCATTTCCTCAATGTGATTTGTTCACTTACATCTGTTTACGTTATG	2400
CTCTTCTCAGGGGAAATTTATGAGCTGTATGAGCTAGATGCGGAATTAG	2450
GTCACAAGGAGCTTGACTCGAAGATGACCATGAGTCGTGTGGGAAAATGC	2500
AGACAGGTAAATTAGTTGAAACAACTGGCCTGCCTTGAATTATTGTGTCTA	
TARRITTTGACACCACCITTTGTTTCAGGTTCGTATCTCTGAAAGTGGGA	
TAGATGAGGCAGTGCAAAAGCTATTAGCTCGTGGGTAAGGGAACCATCAT	- 44.0
actitatggaattegtitaetgetactteggetaggatitaagaaatgga	
AATCACTTCAAGCATCATTAGTTAGGATCCTGAGAACTCAGGATGTTTTC	
TTATTEGTTATATAAAAGTCTTTTCRTCRAGGAGTAACAAAAACTT	
GCACAATATTTGTGTGCTCACTGGCAAGGCATATATACCCAGCTAACCTT	
TGCTAGTTCACTGTAGTAACAGTTACGGATAATATATGTTTACTTGTATG	
TGGTACCCTCATTTTGTCTCTCATGGAGGCTTTGAAGCCTTGTGTTGAAA	
CTGGATAGTTACATATGCTTCCAACAGAAACTAGCATGCAGATTCATATC	,
CTTTCTTATTCTACTARTATGTATGTATCACACACTCGTTGTTTTTTTCA	
AGATATAAAGTTGGACCAATECAGCAGCTAGAAACATCTGACCAAGCAA	
AGCCAGAGGCCAGTACTGTAAGTTTTCGTAGATAGGCCAAAGAGAGAG	
TTGEAGACTGTTTTTGATCATTTCTTTTTTGTACATTACTTTCATGCT	3
AJJAKSSTTAATABASTKTYKSSSSSSSTSTANSSSTAKCTSAATTAKT	
GCTAGTTCAGGTATTAACTCCATCAACAGCAAGCGAGGAAAACATCGCG	
CTGATGCCGTCCATCTTCTTGCTATAAAAGAGGTTTGTTATTTACTTAT	
TATOTTATEATGTGAGTCATCATCATCAAAAAAAAAAAAAAAAA	
TACCARTCTTCCATCARGCTGTGTAAAGGATTTCCAATTAGAAAATCAT	3400
ATTTGATCCTTTGTTTTATATCCAASAGSTCCCTTGAAAAGATCTGTT	3450
AAGATTETTTGCACTTGAAAAATTUAATETTTTTAAGTGAATEEEETAC	T 3500
TTCTTACAATGATCATAGTCTGCAATTGCATGTCAAGTAATATCATTCC	3550
*STTACTSCATECCCCTCTTICTTAATGACCATTGTCTAIGTTGTSTTI	3600
TCTCGTGTGCTGGAGAAAATGATAGCTGATCCAAGCTCIACATTATCAT	
ATTAAGTAGCTGCTCAGGAATTGCCTTTGGTTACATTGCCTAATGGT77	7B 3700
ATGTCAATTTTTCTTCTGAATCTTTATTTTAGATCAAARTGGAGCTRCA	3750
ARGTGTTCARCTGTGTATGGATTTGCTTTTGTTGACTGTCCTGCCTTGA	4G 3BB0
GTTTTGGGTTGGCTCCATCAGCGATGATGCATGATGTCCTGCTCTTGG	AG 3850
CGTTATTGATGCAGGTAAGCAAGTGTATTCTGTATCTTATGTGTACCAT	rg 3900
TGACTTCCTGTGCATATATTTGGGTTGCAGGAACTAATTCTGAATCAC	CA 3950
TTTGGTATGTTTTTTCCAGGTTTCTCCAPAGGAAGTGTTATATGACAG	TA 1000
AAGGYAAACTECTTGTATCGCCAGTTGTTTTGTTAAACAGAATTTAAG	GT 4050
AAA TGACACTGGTTAATTTAAAGTGCATACATGTTGAAATATTGCAGG	GC 4100
TATCARGAGARGCACAAAAGGCTETARGGAAATATACGTTGACAGGTA	CC 4150
ATTTCAGTAGGCAAGCTAACTGACAATTTAACCGCTCACCGAATGATA	GG 4200
TETETTARACATTGETAATGTAGATGATGTTTATGTTTCAATCTAATA	GG 4250
GTCTACGGCGGTACAGTTGGCTCCAGTACCACAAGTAATGGGGGATAC	AG 4300
ATGCTGCTGGAGTTAGAAATATAATAGAATCTAACGGATACTTTAAAG	GT 4350
TCTTCTGAATCATGGAACTGTGCTGTTGATGGTCTAAATGAATG	GT 1400
101101000000000000000000000000000000000	

Figure 11 (Continued)

TGCCCTTAGTGCTCTTGGAGAGETAATTAATCATCTGTCTAGGCTAAAGG	4450
TOTOTOGGCTTGTTTAGTTTTTGCTTTTCACAAATTAAGCAAAGGAACTT	4500
TTCATAACTTACAGTTTCTATCTACTTGCAGCTAGAAGATGTACTTAAGC	4550
ATGGGGATATTTTTCCATACCAAGTTTACAGGGGTTGTCTCAGAATTGAT	4 600
GCCCASACGATGGTAAATCTTGAGATATTTAACAATAGCTGTGATGGTGG	4650
TCCTTCAGGCARGTGCATATTTCTTTTTTGATAACTTCAACTAGAGGGCA	4700
GACATAGAAGGAAAAATTCTAATACTTCGTACGGATCTCCAGTAAGTA	4750
AGCCGATTTTGTTTACCTATGTAGGGACCTTGTACAAATATCTTGATAA	4900
CTGTGTTAGTCCAACTGGTAAGCGACTCTTAAGGAATTGGATCTGCCATC	4850
CACTCAAGATGTAGAAAGCATCAATAAACGGCTTGATGTAGTTGAAGAA	4900
CACTCADAGATGTAGAAAGCATCAATAAACOOCTTCATCACTATCTCCACAA	4950
TTCACGGCAPACTCAGRAAGTNTGCAAATCNCTGGCCAGTATCTCCACAA	5000
ACTTCCAGRETTAGAARGACTBCTCGGACGCATCAAGTCTRGCCTTCGAT	5050
CATCAGCCTCTGTGCCTCTTCTTGGGGAAAAAGTCCTGAAACAA	5100
CGAGTAAGTATCAATCACAAGTTTTCTGAGTAATGCCTTCCATGAGTAGT	5150
ATAGGACTAAAACATTACGGGTCTACCTAAAGACTCTTCTCCTTCTTTTG	5200
CAATGTCTGGTTATTCATTACATTTCTCTTAACTTATTGCATTGCAGGTT	5250
AAAGCATTTGGGCAAATTGTGAAAGGGTTCAGAAGTGGAATTGATCTGTT	5300
GTTGGCTCTACACAAGGAATCAAATATGATGAGTTTGCTTTATAAACTCT	5350
GTAAACTTCCTATATTAGTAGGAAAAAGCCGGXCTAGAGTTATTTCTTTCT	5400
CAATTCGAAGCAGCCATAGATAGCGACTTTCGAAATTATCAGGTCCCCAT	5450
CTATCTTTCATACTTTACAACAAAATGTCTGTCACTACTCAAAAGCAATGC	5500
ATATGGCTTAGATCTCAACTCACACCCCGAGGATCCTAAAGGGATTTGCT	5550
TTTTATTCCTAATGTTTTBGATGGTTTGATTTATTTCTAACTTGAACTT	5600.
ATTAATETTGTACEAGAACCAAGATGTGACAGATGAAAAACGCTGAAACTC	5650.
TCACAATACTTATCGAACTTTTATCGAAAGAGCAACTCAATGGTCTGAG	5700
GTCATTERCACCATAAGCTGECTAGATGTOCTGAGATETTTTGEAATCGC	5750-
AGCARGTCTCTCTGCTGGAAGCATGGCCAGGCCTGTTATTTTTCCCGAAT	5800
CAGAAGCINCAGATCAGAATCAGAAAACAAAAGGGCCAATACTINAAATC	5850
CAAGGACTATGGCATCCATTTGCAGTTBCAGCCGATGGTCAATTGCCTGT	5900
TCCGRATGATATACTCCTTGGCGAGGGTAGAAGAAGCAGTGGCAGCATTC	5950
ATECTCGGTCATTGTTACTGACEGGACCAAACATGGGCGGAAAATCAACT	6000
CTTCTTCGTGCAACATGTCTGGCCGTTATCTTTGCCCAAGTTTGTATACT	6050
DGTTAGATAATTACTCTATTCTTTGCAAYCAGTTCTTCAACATGAATAAT	. 6100
RAATTCTGTTTTCTGTCTGCRGCTTGGCTACGTGCCCTGTGAGTCTT	6150
GCGAARTCTCCCTCGTGGATACTATCTTCACAAGGCTTGGCGCATCTGAT	6200
AGAATCATGACAGGAGAGAGTAAGTTTTCTTCTCAAAATACCAATTCCTC	6250
Gractatttactcagatittgtctgattggacaaggtggttttgcttttt	•
TTTAGGTACCTTTTTGGTAGAATGCACTGAGACAGCGTCAGTTCTTCAGA	6300
ATGCAACTCAGGATTCACTAGTAATCCTTGACGAACTEGGCAGAACT	6350
AGTACTTTCGATGGATACGCCATTGCATACTCGGTAACCTGCTCTTCTCC	6400
TTCAACTTATACTTGTTGATCAACAAAAACATGCAATTCATTTTGCTGAA	64\$0
ACTTATTGATTTATATCAGGTTTTTCGTCACCTGGTAGAGAAGTTCAAT	6500
GTOGGATGCTCTTTGCAACACATTACCACCCTCTCACCAAGGAATTCGCG	6550
TCTCROCCACETCTCACCTCGAARCACATGGCTTGCCCATTCAAATCAAG	6600

ATCTGATTATCAACCACGTGGTTGTGATCAAGACCTAGTGTTCTTGTACC	6650
GTTTAACCGAGGGAGCTTGTCCTGAGAGCTACGGACTTCAAGTGGCACTC	6700
ATGGCTGGAATACCAARCCAAGTGGTTGAAACAGCATCAGGTGCTGCTCA	6750
agccatgargagatcaattggggaaaacttcaagtcaagt	6800
CTGAGTTCTCAAGTCTGCATGAAGACTGGCTCAAGTCATTGGTGGGTATT	6850
TCTCGAGTCGCCCACAACAATGCCCCCATTGGCGAAGATGACTACGACAC	6900
TITGTTTTGCTTATGGCATGAGATCAAATCCTCTTACTGTGTTCCCAAAT	695 0
AAATGGCTATGACATAACACTATCTGAAGCTCGTTAAGTCTTTTGCTTCT	7000
CTGATGTTTATTCCTCTTAAAAAATGCTTATATATCAAAAARTTGTTTCC	7050
TCGATTATAACAAGATTATATATGTATCTGTCGGTT7AGCTATGGTATAT	7100
AATATATGTATGTTCATGAGATTGGTCAAGAGAAATACTCACAAACAGTA	715D
TATTAAGAAGGAAATATGTTTATGCATTAATTTAAGTTTCAAGATAAACT	7200
GCAAATRACCTCGRCTAAAGTTGCAAAGACCRAACACAAATTACAAAACT	7250
TATABBACTTARGTTCTGAATTCCCTARAACAAAAAAAAAAAAAAA	7300
TATTTTGTTGCATCTACAARCAACACAAACCTACATAGTTTATAACTTAC	7350
TCATCACTGAGAT TAACATCAGAATCAT TCTCCATTTCTTCATCTTCACT	7400
CTCATCATCACCACCACCATGATGATTCTCCTCCTCTCACGTAACC	745D
TAGGAATCTCACTCTGAGGTGTATCAACAATCTGGTTCTTCTGGAAGTCC	7500
ANATOTOTOTGANANTCAGCTCTCATCTTCTCCAACTCCTTCATTTGCTC	7550
TTTCTTACTCTTCTCCATCTTCTCATAAACCTTCCCAAACCTCTCAACAG	7600
antecgcearcatettracgaageagegteattareettetteeteeg	7650
TACTCAACCTCATCCTCATCCTCCTCCTCTTCAGAATCACCAGGACT	7700
ATCCATCATCATCAAACCCRTTAGACTTATCTAAATAAACCTTAGTGT	7750
TCATAAACACAAACTCACCTGAATCAACACCACAAGCTAAACCTAAATCC	7800
GACTTGGGGGAAACACAAAGCAACATATCCAACTTATTGAAAAACGACCA	7850
TTTACTTGAACCTAAACCTGATTTCTCAACCTTAATCTTCTCTTTTCTAT	7900
ACTTCCTCTTCAAGTCATCATCATCTCCTACATTGCGTCTCAGATTTC	7950
TECATCETTAGETECTCACTCACTTTCTCAGCTACTTCATTCCAATCCTC	800B
GTTCCTCARACTCCTTCTACCCAATTGCRAAAACCTATCTCCCCAAACTT	9050
CAAGCAACACAA	BD62

Figure 11 (Continued)

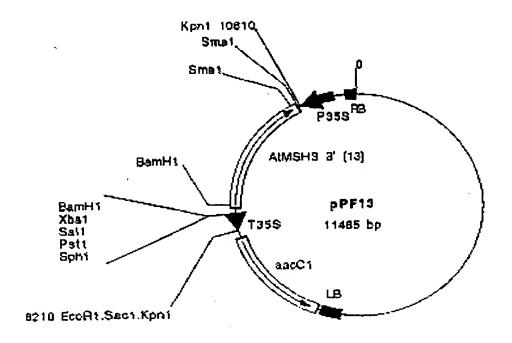


Figure 12

Commente/References: AtMSH3 3' side antisense : AtMSH3 3' (19 = 2104bp) from pUC18/13 Sal1/Sal1/T4 Into pCW184 BamH1/T4 In Agrobacterium LBA44O4

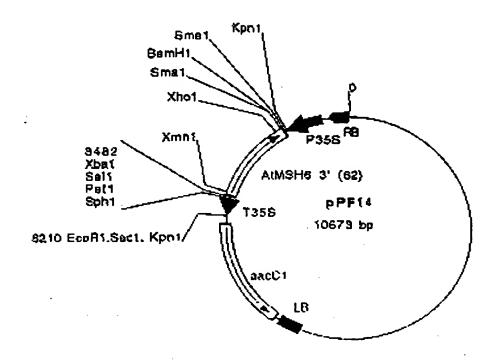


Figure 13

Comments/References: AtMSH6 (S8) 3' side antisens : 62 Sal1/Sal1/T4 (1379bp) into pCW164 BamH1/T4

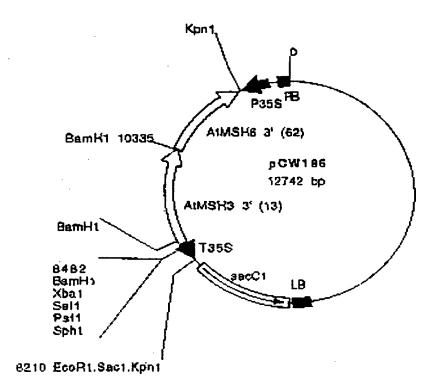


Figure 14

Comments/References: AtMSH6 3'/AtMSH3 3' antisense : AtMSH6 (SB) 3' side (62=1379bp) Sal1/Sat1/T4 Into pPF13 (pCW184 AtMSH3 (S5) 3' side (13=2104) antisens)/Smal. in LBA4404

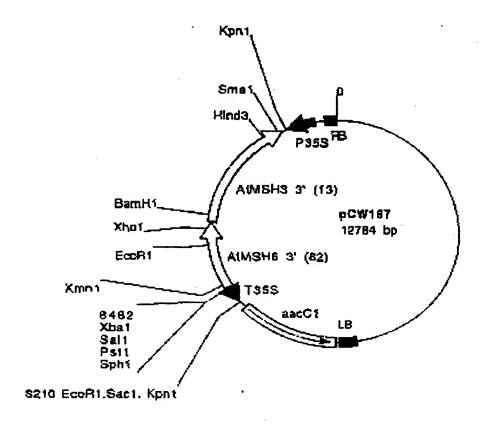


Figure 15

Comments/References: AtMSH3 3'/AtMSH6 3' antisens (0): AtMSH3 (35) 3' eide (13=2104bp) Sat1/Sst1/T4 into pPF14 (AtMSH6 (35) 3'eide (62=1379bp) antisense into pCW164)/Sma1. In LBA4404

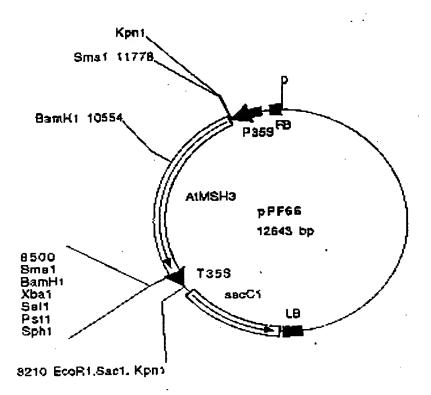


Figure 16

Comments/References: AtMSH3 (SB) complete, sease prientation : pPF26 (3942bp) Email Into pCW164 Small

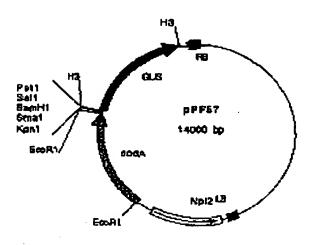


Figure 17

Commente/References: pPZP111 with code Ecofit casestle in Ecofit ails and Hinda GUS casestle in Hinda site. Kenfl. All genes under Prompter/reminator 353

Figure 18

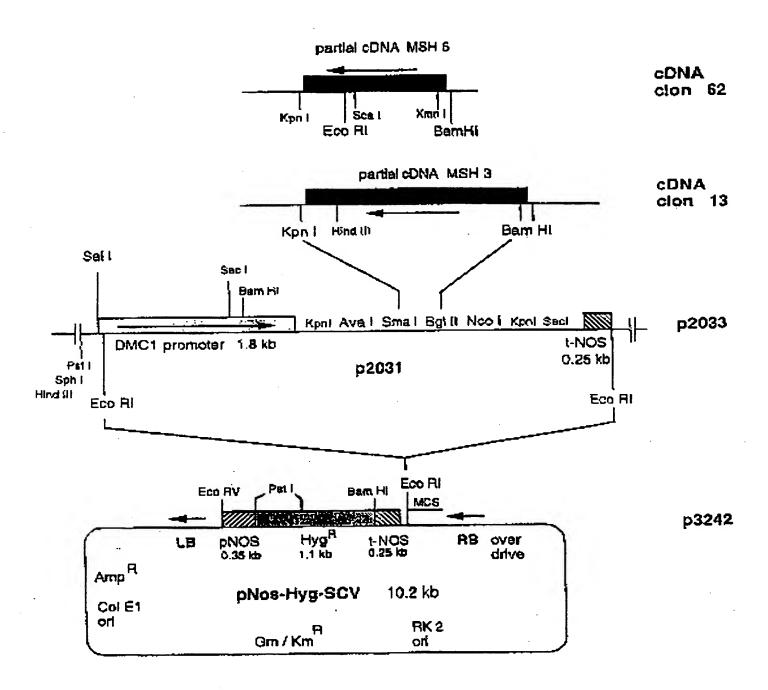
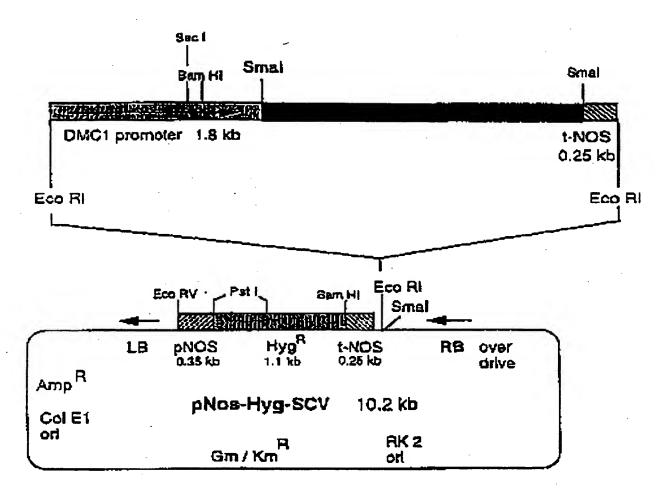


Figure 19

p3243



1

SEQUENCE LISTING

```
Rhone-Poulenc Agro; Betzner, Andreas Stefan; Doutriaux,
<110>
             Marte-Pascale; Freyssinet, Georges; Perez, Pascual.
<120>
             Methods for obtaining plant varieties
             395498C
<130>
<150>
             PO9745
             1997-10-10
<151>
             98
<1605
<210>
             33
<211>
<212>
             DIVA
<213>
             Artificial sequence.
<220>
             modified_base
<221>
c222%
             11
<223>
             L
<220>
<221>
             modified_base
<222×
             14
<223×
             I
<220>
<221>
             modified base
<222>
             17
<223>
             I
<220>
             Degenerate oligonucleotides UPMU used to isolate AtMSH3 and
<223>
             AtMSH6.
c300>
             Reenan and Kolodner
<301>
<302>
             Genetica
<303>
             132
<306>
             963-973
<307>
             1992
<400>
                                                                         23
ctggatccac nggnecnaay atg
```

VSDOCID: <WO__9919492A2TI_>

<210><211>

<212>

23 DNA

```
<213>
             Artificial sequence
<220×
<221>
             modified_base
<222>
             15
<223>
             1
<220>
             modified_base
<221>
             18
<222>
<223>
             I
<220>
             Degenerate oligonucleotides DOMU used to isolate AtMSH3 and
<223>
             Atmsh6.
<300>
             Reenan and Kolodner
<301>
<3025
             Genetics
<303>
             132
<306>
             963-973
<307×
             1992
<400>
             2
                                                                         23
etggateert ortgogtore raa
<210>
<211>
              24
<212>
             DMA
             Artificial sequence
<213>
<320×
              MSR3 specific primer 636 for PCR using cDNA of Arabidopsis
<223>
              thaliana ecotype Columbia
<400>
              3
                                                                         24
tgctagtgcc tcttgcaage tcat
<210>
              4
              27
<211>
<212>
              DNA
<213>
              Artificial sequence
<220>
              Primer AP1 for PCR using cDNA of Arabidopsis thaliana ecotype
<223>
              Columbia containing adapter sequences ligated to both its
              enda
<400>
```

•	
ccatcctast a	coacteact ataggec 27
03.0	5
<210>	33 2
<211>	DNA
<212>	
<213>	Artificial sequence
.nan.	
<220>	Primer AP2 for PCR using cDNA of Arabidopsis thalians ecotype
<223>	Columbia containing adapter sequences ligated to both its
	ends
	÷1:05
<400×	5
1.00	-
actractata (gggetegage ggc 23
<210>	6
<211×	30
<212>	DNA
<2135	Artificial sequence
<22 0 >	
<233>	MSH3 specific primer S525 for PCR using cDNA of Arabidopsis
	thaliana ecotype Columbia
<400>	6
	unraratano aattsectte 30
aggittetgat	tatgtgtgae getttaetta 30
<210>	7
<211>	29
<212>	DOVA
<213>	Artificial sequence
40257	
<220>	
<223>	MEH3 specific primer SS1 for PCR using cDNA of Arabidopsis
	thaliana ecotype Columbia
<400>	7
ggatogggta	ctgggttttg agtgtgagg
<210>	· · · · · · · · · · · · · · · · · · ·
<211>	24
<212>	DNA
<213>	Arcificial sequence
<220>	
<223>	MSH3 specific primer 635 for PCR using cDNA of Arabidopsis
	thaliana ecotype Columbia

<400>	3
gracgtgett	gatggtgttt teac 24
<210><211><212><213>	9 20 DNA Artificíal sequence
<220> <223>	MSH3 specific primer 8523 for PCR using cONA of Arabidopsis thaliana ecotype Columbia
<400>	9
tcagacagta	tecageatgg cagaagta 28
<2105	10
<211>	33
c212>	DNA
<213>	Artificial sequence
<220>	
<223>	MSH3 specific primer 185 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
₹400>	10
atcccgggat	gggcsagcaa aagcagraga cga 33
<210>	11
<211>	27 .
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	MSH3 specific primer SS3 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	11
gacaaagagc	gaaatgagge recttgg 27
<210>	12
<211>	1250
<212>	DNA
<213>	Arabidopsis thaliana ecotype Columbia
<223>	Clone 52

<400> 12

cccggg&cgg	gcaagcaaaa	gcagcagacg	acttotogtt	tettegetee	Caaacccasa	60
terregarte	асдавссдаа	teeggtagee	gaatcatçag	cacegeeace	gaagetatcc	120
gccactgtat	ccttctctcc	ttccaagegt	aagcttctct	ccgaccacet	càccaccaca	180
teacceaaaa	agectaaact	ctotootoae	octcaasacc	cagtacccga	tcccaattta	240
caccaaagat	ctcccagag	atttctggsa	ccctcqcegg	aggaatatgt	recegaaacg	300
ccatcatcga	ggaaatacac	accattggaa	cagcaagtgg	tggagctaaa	gagcaagtac	360
ccagatgtgg	ttttgatggt	ggaagttggt	tacaggtaca	gattettegg	sgaagacgcg	420
gagatogcag	cacgcgtgtt	gggtatttac	geteatatgg	atcaceattt	cstgscggcg	480
agtgtgccaa	catttegatt	gaaunuddau	gtgagaagac	tggtgaatgc	aggatacaag	540
atcggtgcag	cgaagcagac	tgaaactgca	gccattaagt	cccacggcgc	aaaccggacc	600
ggccottttt	tcoggggact	gtjaggagtig	tataccaaag	ceaegettga	agcggctgag	660
gatataagtg	grągrigtąg	tggtgaagaa	ggttttggtt	cacagagtaa	tttcctggtt	720
tgtgt tgtgg	atgagagagt	taaguoggag	acattaggct	gtggtattga	aatgagtttt	790
gatgttagag	toggtgttgt	tggegttgaa	atttegacag	gtgaagetgt	tcatgaageg	. 840
ttcaatgata	actteatgag	aagtggatta	gaggetgtga	ttttgagett	gtcaccaget	900
gagotgttgo	ttggccagcc	totttcacaa	casactigaga	agttittggt	ggcacatgct	960
ggacctaret	caaacgttcg	agtggaacgt	gcctcactgg	attgtttcag	caecggtaat	1020
gcagtagatg	aggttatttc	attatgtgaa	aasatcagcg	caggtaactt	agaagatgat	1080
aaagaaatga	ağır gəgəgə	tgctgsaaaa	ggaatgtett	gcttgacagt	tcatacaatt	114D
atgaacatgo	cacatrigae	tgttcaagcc	ctegecetaa	egttttgcca	teteaaacag	1200
tttggatttg	aəaggateet	ttaccaaggg	gestsættts	getetttgte		1250

<210> 13 <211> 34 <212> DNA <213> Artificial asquance

<220>

<2235 MSH3 specific primer 255 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400>	13					
attongggtt	aaaatgaaca	agroggettt	agtc			34
<210>	14					
	27					
<211>					•	
<212><213>	DNA Arsificia	l sequence				
12137	MIUILIUAG	r seduence				
<220>						
<223>	MSH3 spec	ific orimer	S52 for PC	R usino ∈DN	A of Arabido	CRIR
;=		ecotype Col				
<4DO>	14					
gccacatetg	actgttcasg	ccetege			•	27
	_					
<210>	15					•
<211>	2110					
<212>	DNA					
<213>		saniind ai	ecotype Co	11m019		•
e223>	Clone 13					
<400>	1.5					
gccacatetg	actgttcaag	ccetegeeet	aacgttttgc	catotosaac	agtttggatt	64
tgaaaggatc	ctttaccaag	gggcctcatt	tcgctctttg	teaagtiaca	çagagatgac	121
CEEEEcagec	aatactccgc	aacagttgga	ggttgtgååå	aataattcag	atggatcgga	180
atotggetce	ttattccata	atatgaatca	cacacttaca	gratatggtt	ceaggettet	240
tagacactgg	gugacucate	ctctatgcga	tagaaatttg	statetgete	ggcttgatgc	30
tgtttetgag	atttetgett	gcatgggatc	teatagttet	toccagetca	gcagtgagtt	360
ggttgaagaa	ggttetgaga	gagcááttgt	atcacctgag	tittatctcg	tgatataata	42
agtetigaca	gctatgtcta	gatcatctga	tattcamcgt	ggaataacaa	gaatetttea	48
toggactget	asagecaeaġ	agttcattgc	agttatggaa	gctattttac	tigcggggaa	54
gcaaattrag	cggcttggca	taaagcaaga	ctctgaaatg	aggagtatgd	astotgoaac	6 D
tgtgcgafct	actettttga	gaaaattgat	ttctgttatt	tcatcccetg	ttgtggttga	66
caatgccgga	aaacttctct	ctgccctaaa	tanggangeg	gctgttcgag	gtgacttgct	72
cgacatarta	atcacttcca	gegaecaatt	tcctgagctC	gctgaagctc	gccaagcagt	78
******	30000000000	tagatt cat a	art aportés	******	4000 DOC+ 1 F	94

tcgaaatttg	gaatttcttc	aagtgtcggg	gatcacacat	ttgatagagc	tgcccgttga	900
ticcaaggic	cctatgaatt	<u>dååtasse</u>	aaatagcacç	aagaagacta	thegatatea	960
cccccagaa	acagcagetg	gcttggatga	gctagczeta	gcaactgaac	arcttgccat	1020
tgtgaaccga	gette ştg gg	ategtttcct	caagagtttc	agtagatact	acacagattt	1080
taaggetgee	gttcaagctc	ttgctgcact	ggactgtttg	cactcccttt	caactctatc	1140
tagaaacaag	aactatgtcc	gtecegagtt	tgtggatgac	tgtgaaccag	ttgagataaa	1200
catacagtct	ggtcgtcatc	ctgtactgga	gactatatta	caagataact	tegteccaaa	1260
tgacacaatt	regratgeag	aaggggaata	ttgccatatt	atcaccggae	ctaacatggg	1320
аддавададс	tgctatatcc	gtcaagttgc	tttaattt.cc	ataatggctc	aggttgg tt c	1380
ctttgtacca	gcgtcattcg	ccsagetgea	cgtgcttgat	ggtgttttca	cteggatggg	1440
tgcttragac	agtatecage	atggcagaag	tacciticta	gaagsattaa	gtgaagegte	1,500
agacataatc	agaacotgut	cttctcgttc	gcttgttata	ttagaugago	trggaagagg),560
ractageaca	<u> </u>	nagodatugo	ctatgcaaca	ttacagcatc	toctageaga	1620
aaagagatgt	ttggttettt	ttgtczcgca	ttaccctgaa	atagetgaga	tcagtaacgg	1690
attereaggt	cctgttggga	cataccatgt	ctcgtatctg	acattgcaga	aggatəaagg	1740
cagttatgat	catgatgatg	cgacctacct	alataagott	gtgcgtggto	tttgcagcag	.1900
gagetttggt	tttaaggttg	oteagettge	ccagetacct	ecatcatgta	tacgtegage	1860
catttcaatg	gctgcaaaat	tggzagetga	ggcacgcgca	agagagagaa	atacacgcat	1920
gggagaacca	gaaggacatg	aagaaccgag	aggegeagaa	gaatctattt	cggctctagg	1980
tgacttgttt	gcagacctga	aatttgctct	ctctgaagag	gacccttgga	aagcattcga	2040
gtttttaaag	catgettgga	agattgetgg	caaaatcaga	ctaaaaccaa	cttgttcstt	2100
ttgaceeggg	·				·	2110

<210>	16
<211>	29
<213>	DNA
<213>	Artificial seguence
<220>	
<223>	MSH3 specific primer S\$1 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400>

ädarcäääre	ctgggttttg agtgtgagg	29
<210>	17	
<211>	30	
<212>	DNA	
<213>	Artificial seguence	
1220		
<220>		
<223>	MSH3 specific primer S525 for PCR using cDNA of Arabid	loosis
·	thaliana erotype Columbia	
<400>	17	
aggitetgat	tatgtgtgac gctttactta	3 0
<210>	18	
<211>	3522	
<212>	DNA	
c213>	Arabidopsis thaliana ecotype Columbia	
<220>		
<221>	CD\$	
<272>	(100)(3342)	
<223>	AtMSR3 full-length cDNA and deduced sequence of the en	:೧೦ರೇಗ
	polypeptide	
<400>	18	
ccsaagasag	egegegaaaa tiggeaacee aagtiegrea tagreargae catgatette	60
catttetett	aaacggagga gattacgaat aaagcaact	99
		4 4 7
	g can and can can ace att tot cot tto tto got oce and	147
	a Gln Lya Gln Gln Thr Ile Ser Arg Phe Phe Ala Pro Lya	
1	5 10 15	
666 373 to	G 007 384 638 633 685 834 888 643 886 888	195
	c ong act ose gas ong ast ong gts gon gas tok tok ack	473
FIG DAS SO	r Pro Thr His Glu Pro Asn Pro Val Ala Glu Ser Ser Thr 20 25 30	
	20 25 30	
207 CAR DO		243
	g aag ata too goo act gta too tto tot oot too aag ogt o Lys Ile Ser Ala Thr Val Ser Phe Ser Pro Ser Lys Arg	£ 12 -3
3	· · · · · · · · · · · · · · · · · · ·	
•	J 10 15	
aad off of	C too gae cad ofe goe goe gog toa coe aaa aag cot aaa	291
-	u Ser Asp His Leu Ala Ala Ala Ser Pro Lys Lys Pro Lys	
50	55 60	
20	33 6 V	
	t cac act caa aac coa gta coe gat coe aat tta cac caa	339
	o His Thr Gln Asn Pro Val Pro Asp Pro Asn Leu His Gln	~ ~
65	70 75 80	
		

_							gec Pro		_	_					387
							aca Thr 105		-						435
							gtg. Val							•	483
							ysb gsc					-	_		531
	•			-		-	cac Nis						_		579
			•	_			gtg Val	_	_	_			_		627
							act Thr 185								675
							uut Phe								723
			-	_			gct Ala								771
							cag Gln	_							819
					ràs		aca Thr								867
-		_	_	_	_		gtt Val 265		-	_		_			915
-	-	-		_			gat Asp			_	_	•			963

WO 99/19492 PCT/EP98/06977

											gly ggc		1011
											gga Gly		i ņ 59
											agc Ser 335		1107
											agc Ser		1155
											gaa Glu		1203
 _	_	_		•				•	_		cat %is	_	1251
	_		-		•		-			_	ttt Phe		1299
											tca Ser 415		1347
					_			_	_	_	gag Glu	•	1395
											cat His		1443
											cac Xís		1492
											ctt Leu		1539
						atg Met					tcc Sex 495	_	1587

											att 1le 510		1635
							_	-			atg Met		1683
	_										yrg cgg		1731
								-			ctt Leu	 	1779
											atg Met		1627
											ttg Leu 590		1875
			-		_						ctt Leu		1923
											Asp		1971
											ege		2019
_	_	_	 	-		•	_	Ser			tca Ser	 	2067
											teg Sør 670		2115
	_		 _		_	qeA		•	•		atg Kis		2163
										His	Pro		2211

ata Ile 705	Val	gct Ala	gjå ååc	Etg	gat Asp 710	gag Glu	cta Leu	gct Ala	cta Leu	gca Ala 715	act Thr	gaa Glu	cat Ris	peir ctr	900 Ala 720	2259
Ile	Val	ABN	Arg	Ala 725	Ser	Тхр	Уsb	Ser	Phe 730	Leu	Lys	ser	Phe	agt Ser 735	Arg	2307
Tyr	Tyr	Thr	740	Phe	្រកន	Ala	Ala	Val 745	Gln	Ala	Leu	Ala	Ala 750	ctg Leu	ysb	2355
Cys	ГБЛ	Жів 755	Ser	Leu	Ser	Thr	1-eu 750	Ser	Arg	Aen	Lys	Asn 765	Tyr	gtc Val	Arg	2903
Pro	770	Phe	Val	Asp	Asp	Ր ሃ ։ 775	Gla	Pro	Val	Gl'n	lie 780	Авл	lle	cag Gln	Ser	2451
Gly 785	Arg	aiH	Pro	Val	190 790	G1:a	7	ile	ГЅЛ	Gln 795	ĄSÞ	usv	Phe	gtc Val	900	2499
Asn	Авр	Thr	Ile	Leu 805	Ris	Ala	Glu	ülу	Glu 810	Tyr	Суs	Gln	Ile	ate 11e 915	Thr	2547
Oly	Pro	ask	Met 620	Gly	Gly	Lys	Ser	C'ys 625	Tyr	Ile	Arg	Gln	Val 830	y¢t Ala	Leu	2595
Ile	Ser	11e 835	Met	Ala	Gln	Val	G1y 640	Ser	Pre	Val	Pro	Ala 845	Ser	ttc Phe	Ala	2643
Lys	ctg Leu 850	cac His	gtg Val	ctt Leu	gat Asp	ggt 61y 855	gtt Val	tte Phe	açt Thr	yrg	atg Met 860	ggt Gly	gct Ala	tca Ser	gac Asp	2691
agt Ser 965	lle	cag Gln	cat His	ejà aac	aga Arg 870	agt Ser	acc Thy	ttt Phe	cta Leu	gaa Glu 675	gaa Glu	tta Leu	agt Ser	gaa Glu	gcg Ala 880	2739
tça Ser	cac Hia	ata Ile	atc [le	aga Arg 885	acc Thr	сув Сув	tct Ser	tct Ser	egt Arg B90	tcg Ser	ctt L e u	gtt Val	ata Ile	tta Leu 895	gat Asp	2787
gag Glu	ctt Leu	G1y gga	ano yra gas	Gly ggc	act Thr	agc Ser	aca Thr	cac Kis 905	gac Asp	Gly	gta Val	gcc Ala	att Ile 910	gcc Ala	tat Tyr	2835

gca aca tts cag cat ctc cts gca gaa aag aga tgt ttg gtt ctt tt Ala Thr Leu Gln His Lou L u Ala Glu Lys Arg Cys Lou Val Leu Pho 915 920 925	t 2883 e .
gtc acg cat tac ect gam ata get gag atc agt aac gga ttc eca gg: Val Thr Ria Tyr Pro Glu Ile Ala Glu Ile Ser Aan Gly Phe Pro Gly 930 935 940	
tet get ggg aca tac cat gto tog tat ong aca trg dag dag gat aa. Ser Val Gly Thr Tyr Ris Val Ser Tyr Leu Thr Leu Glo Lys Asp Ly: 945 950 955 96	s
ggc agt tat gat cat gat gat gtg acc tac cta tat and ctt gtg cg Gly Ser Tyr Amp Him Amp Amp Val Thr Tyr Leu Tyr Lym Leu Val Ar 965 970 975	
ggt ctt tgc age agg age ttt ggt ttt aag gtt get cag ett gee ca Gly Leu Cys Ser Arg Ser Phe Gly Phe Lys Val Ala Glo Leu Ala Gl 980 985 990	g 3075 n
ata out one hea tigt ata out our gor and toe and got gos and the Pro Pro Ser Cys lie Arg Arg Ala The Ser Met Ala Ala Lys Le 1995 1000 1005	
gaa got gag gts ogt goa aga gag aga aan aca ogo atg gga gaa oo Glu Ala Glu Val Arg Ala Arg Glu Arg Asn Thr Arg Met Gly Glu Pr 1818 1820	a 3171 o
gas ggs cat gsa gas ceg aga gge ges gas gas tet att teg get et Glu Gly His Glu Glu Pro Arg Gly Ala Glu Glu Ser Ile Ser Ala Le 1025 1030 1035 104	บ
ggt gad tig tit goa gad dig aas hit gon die ien gaa gag gad de Gly Asp Leu Phe Ala Asp Leu Lys Phe Ala Lou Sor Glu Glu Asp Pr 1045 1050 1055	et 3267 o
tgg aaa goa tro gag ttt tta aag cat got tgg aag att got ggc aa Trp Lys Ala Phe Glu Phe Leu Lys Ris Ala Trp Lys Ile Ala Gly Ly 1060 1065 1070	
atc aga cta ama cca act tgt tca ttt tgatttamtc ttamcattat lle Arg Leu Lys Pro Thr Cys Ser Phe 1075 1080	3362
agraactyca agytettyat catetyttag ttycytacta acttatytyt attayt	1taa 3422
caagaaaaga ganttagaga gatggattot aatcoggtgt tgcagtacat otttto	cca 3482
ccccatada aaqaadaaaa aaaaaaaaa aaaaaaaaaa	3522

<210> 19 <211> 10B1 <212> PRT

<21)					psis Tide			ાત ૧૯	otyj	ise Co	չ և տան	oia			
< 4 🗆 () »		19												
Met l	gly	Lys	Gln	Lув Ş	Oln	alo	Thr	Ile	Ser 10	Arg	Рhе	Phe	Ala	Pro 15	Lys
Pr	Lys	Ser	20 20	Thr	Kia	Сĵя	Pro	Asn 25	Pro	val	Ala	ĠĴи	Ser 30	Ser	Tha
· Pro	Pro	Pro 35	ьуs	lle	Ser	Ala	Thr 40	ral	ser	₽уе	Ser	Pro 45	ser	Lya	Arç
Lys	Նզս 50	Leu	Sor	Азр	His	Leu 55	Ala	Ala	Ala	ser	60 820	Lys	Гув	Pro	Lye
Leu 65	Ser	Pro	Ris	Thr	Gln 70	Aşn	Pro	Va).	B ro	Asp 75	Pro	Asn	ľeu	His	B(
Arg	Phe	Leu	G1v	Arg 85	Pire	Leu	Glu	Pro	Ser 90	Pro	Glu	Gln	тух	Val 95	Pro
Glu	Tier	Ser	5er 100	Sex	λrg	ГÃЗ	Tyr	"hr 105	PTO	ren	Glu	Gin	Gln 110	Val	Val
Glu	Leu	Lув 115	ser	Lys	Тут	Pro	Asp 120	Val	Val	Lev	Met	Val 125	Glu	Val	Ġ15
Tyr	Arg 130	Tyr	Arg	Phe	Phe	Gly 135	Glu	Asp	Ala	Glu	Tie 340	Ala	Ala	Arg	Va!
Leu 145	Cly	lie	Tyr	Ala	His 150	Mec	dev	His	Asn	Phe 155	Met	Thr	Ala	Ser	Va] ·160
Pro	Thr	Phe	Arg	Leu 165	Asn	Phe	おささ	Val	Arg 170	Arg	Leu	Val	aen	Ala 175	ĠΣ
Tyr	Lys	Ile	Gly 180	Val	Val	Lys	Gln	Thr 185	Glu	Thr	Ala	Ala	Ile 190	Гуs	Sex
Kis	Gly	Ala 195	aaa	Arg	Thr	Gly	Pro 200	Phe	Phe	Arg	Gly	Leu 205	Ser	Ala	Let
Tyr	Thr 210	Lys	Ala	Thr	Leu	Glu 215	Ala	Ala	Glu	Asp	11e 220	ser	Gly	Gly	Cye
Gly 225	Gly	Glu	Ģlu	Gly	Phe 230	Gly	Ser	Gln	5er	Asn 235	Phe	Leu	Val-	Сув	Val 240
Lav	Двр	Glu	Arg	Val 245	Lув	Ser	Glu	Thr	Leu 250	Ĝly	Сув	Gly	Ile	Glu 255	Ket

- Ser Phe Asp Val Arg Val Gly Val Val Gly Val Glu Ile Ser Thr Gly 260 265 270
- Glu Val Val Tyr Glu Glu Phe Asn Asp Asn Phe Met Arg Ser Gly Leu 275 280 285
- Glu Ala Val Ile Leu Sor Leu Ser Pro Ala Glu Leu Leu Gly Gln 290 295 300
- Pro Leu Ser Gin Gln Thr Glu Lys Phe Leu Val Ala Met Ala Gly Pro 305 310 315 320
- Thr Ser Asn Val Arg Val Glu Arg Ala Ser Leu Asp Cys Phe Ser Asn 325 330 335
- Gly Ash Ala Val Asp Glu Val Ile Ser Leu Cys Glu Lys Ile Ser Ala 340 345 350
- Gly Ash Len Glo Asp Asp Lys Glo Met Lys Leo Glo Ala Ala Glo Lys 355 360 365
- Gly Met Ber Cya Len Thr Val Eis Thr Ile Met Asn Met Pro Bis Leu 370 375 380
- Thr Val Glm Ala Leu Ala Lou Thr Phe Cys His Leu Lys Glm Phe Gly 385 390 395 400
- Phe Glu Arg Ile Leu Tyr Gln Gly Ala Ser Phe Arg Ser Leu Ser Ser 405 410 415
- Asn Thr Glu Met Thr Leu Ser Ala Asn Thr Leu Gln Gln Leu Glu Val 420 425 430
- Val Lys Asn Asn Ser Asp Gly Ser Glu Scr Gly Ser Leu Phe His Asn 435 440 445
- Met Asn His Thr Leu Thr Val Tyr Gly Ser Arg Leu Leu Arg His Trp 450 460
- Val Thr His Pro Leu Cys Asp Arg Asn Leu Ile Sar Ala Arg Leu Asp 465 470 475 480
- Ala Val Ser Glu Ile Ser Ala Cys Mot Gly Ser His Ser Ser Gln
 485 490 495
- Leu Ser Ber Glu Leu Val Glu Glu Gly Ser Glu Arg Ala Ile Val Ser 500 505 510
- Pro Glu Phe Tyr Leu Val Leu Ser Ser Val Leu Thr Ala Met Ser Arg 515 520 525
- Ser Ser Asp The Glo Arg Gly The Thr Arg The Pha His Arg The Ala 530 540

ъуя 5 45	Ala	ፐከድ	Glu	Phe	11e 550	Ala	Val	Met	Gλů	Ala 555	Ile	Leu	Leu	ala	G13
Lys	Gln	Ile	Gln	Arg 565	Leu	Gly	Ile	Ĺуа	Gln 570	Asp	Ser	Glu	Met	Arg 575	ŝes
Met	Gln	Ser	Ala 580	Thr	Val	Arg	\$er	Thr 505	Leu	Lėu	Arg	Lys	Leu 590	lle	Sei
Val	Ile	Ser 595	Ser	Pro	Val	٧al	Vа) 600	Asp	neA	Ala	Gly	Lys 605	Leu	Leu	561
Ala	leu 610	Asn	Lys	Glu	Ala	Ala 615	Val	Arg	Gly	Asp	Leu 620	Lein	Asp	Ile	Let
11e 625	Thr	Ser	Ser	qzA	630	Phe	Pro	Glu	Lev	Ala 635	Glu	Ala	Arg	Gln	Ala 540
Val	Leu	Val	Ile	Arg 645	61 u	Lys	Leu	Asp	Ser 650	Ser	Ile	Ala	Ser	Phe 655	Arg
l _' ys	Lys	Leu	A) a 660	I)#	Arg	Ash	Leu	665 665	Phe	Leu	Gln	Val	Ser 670	Gly	Ile
Thr	His	Leu 675	Ile	G] Æ	Leu	Pro	Va) 680	Asp	ser	Lys	Val	Pro 685	H1.\$	Asn	Try
٧al	690	Val	Asn	Ser	Thr	Lys 695	Lув	Thr	Ile	Arg	Tyr 700	Hib	Pro	Pro	Glı
Ile 705	Val	Ala	gjà	Leu	Asp 710	GLu	Leu	Ala	Гел	Ala 715	Thr	Glu	Нів	Leu	Ala 720
Ile	Val	Asn	Arg	Ala 725	Ser	Trp	Asp	Ser	Phe 730	Lou	Lys	Ser	Phe	9er 735	Arg
Tyr	Tyr	Thr	Asp 740	Phe	Lys	Ala	Aia	Val 745	Gln	λla	Leu	Ala	Ala 750	Гел	Asg
Cys	Leu	His 75\$	Ser	Leu	Ser	Thr	Leu 760	Ser	Arg	Ast	Lys	Asn 765	Tyr	Val	Arg
Pro	Glu 770	Phe	Val	Λsp	Asp	Сув 775	Glu	Pro	Val	Glu	11e 780	αeA	I <i>l</i> e	Gln	Sei
Gly 785	Arg	нis	Pro	Val	Leu 790	Glu	Thr	Ile	Leu	Gln 795	Asp	ask	Phe	val	Pro
Asn	qaA	Thr	Ile	Leu 805	His	Ala	Glu	GJY	Glu 810	тут	Сув	Gln	Ile	Ile 315	Thi
Gly	Pro	Asn	Met 820	Gly	Gly	Lys	\$e r	Cya 825	Tyr	Ile	Arg	Gln	Val 830		Leu

- The Ser He Met Ala Glm Val Gly Ser Phe Val Pro Ala Ser Phe Ala 835 840 845
- Lys Leu His Val Leu Asp Gly Val Phe Thr Arg Met Gly Ala Ser Asp 850 855 .860
- Ser lle Gln His Gly Arg Ser Thr Phe Leu Glu Glu Leu Ser Glu Ala 865 870 875 880
- Ser His Ile Ile Arg Thr Cys Ser Ser Arg Ser Leu Val Ile Leu Asp 885 890 895
- Glu Leu Gly Arg Gly Thr Ser Thr His Asp Gly Val Ala Ile Ala Tyr 900 905 910
- Ala Thr Leu Gln His Leu Leu Ala Glu Lys Arg Cys Leu Val Leu Phe 915 920 925
- val Thr His Tyr Pro Glu Ile Ala Glu Ile Ser Amn Gly Phe Pro Gly 930 935 940
- Ser Val Gly Thr Tyr His Val Scr Tyr Leu Thr Leu Gln Lys Asp Lyc 945 950 955 960
- Gly Ser Tyr Asp Bis Rep Asp Val Thr Tyr Leu Tyr Lys Leu Val Arg 965 970 975
- Gly Leu Cys Ser Arg Ser Phe Gly Phe Lys Val Ala Gln Leu Ala Gln 980 985 990
- The Pro Pro Ser Cys lie Arg Arg Ala Tie Ser Met Ala Ala Lys Leu 595 1000 1005
- Glu Ala Glu Val Arg Ala Arg Glu Arg Ash Thr Arg Met Gly Glu Pro 1010 1015 1020
- Glu Gly His Glu Glu Pro Arg Gly Ala Glu Glu Ser lle Ser Ala Leu 1025 1030 1035 1040
- Gly Asp Leu Phe Ala Asp Leu Lys Phe Ala Leu Ser Glu Glu Asp Pro 1045 1050 1055
- Trp Lys Ala Phe Glu Phe Leu Lys His Ala Trp Lys Ile Ala Gly Lys 1060 1065 1070
- 11s Arg Leu Lys Pro Thr Cys Ser Phe 1075 1080
- <210> 20
- <211> 24
- <212> DNA
- <213> Artificial sequence

<230>	
<223>	MSH6 specific primer 638 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	20
cctctaccag	gegacgasas accg 24
<210>	21
<211>	28
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Primer S81 for PCR using cDNA of Arabidopsis thaliana ecotype
	Columbia
<400>	21
agtageettt	agcatecct tentreae 28
<210>	23
<211>	30
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	MSH6 specific primer 5823 for PCR veing cDNA of Arabidopsis thaliana ecotype Columbia
<400>	22
gettggcgca	tctaatagaa tcatgacagg 30
<210>	23
<211>	24
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	MSR6 specific primer 637 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	23
gacagegtea	gttetteaga atge 24
-210s	24
<210>	24 25

<212>

DNA

<213>	Artificial sequence	
<230×		
	Mally applific anima 200 for BCD units approach to the biddenic	_
42232	MSH6 specific primer 188 for PCR using cDNA of Arabidopsis	5
	thaliana ecotype Columbia	
<400×	24	
ateregggat	geagegreag agategattt tgt	33
<210>	25	
<211>	27	
<212>	DNA	
<213:	Artificial sequence	
<220>		
c223>	MSH6 specific primer SB3 for PCR using cDNA of Arabidopsi.	В
1-1-1	thaliana ecotype Columbia	_
<400>	25	
	·	
ogotatetat	ggetgetteg aattgag	27
	·	
<210×	26	
<211>	1385	
<212>	DNA	
<213>	Arsbidopsis thalisma ecotype Columbia	
<223>	Clone 43	•
<400>	26	
cccgggargc	agogocagag atogettitg tottlotton aaaaacocac ggoggogact	60
accaaccott	töättteedä edatöetdet väedäedäää äedäegäedä väägeegeede	120
25305353-5	6836666683 484646461 WA6226222 2023962623 WA376667	
tttäätgtga	aggaagggga tgctaaaggc gacgcttctg tacgttttgc tgtttcgmam	180
tetgregatg	aggrragagg aacggatact ccaccggaga aggttccgcg tcgtgtcctg	240
conchosar	ttaagorgge tgaatergre ggtgatgott ogtooctgtt otocaatatt	300
00300033		
atgeataagt	ttgtaaaagt cgatgatcga gattgttctg gagagaggag ccgagaagac	360
		420
gragrecege	tgaatgatto atototatgt atgaaggera atgatgttat tootoaattt	420
cqtcccaata	atggtaaaac tcaagaaaga aaccatgett ttagttteag tgggagaget	480
. <u>-</u>		
gascttagat	cagtagaaga tataggagta gatggcgatg ttcctggtcc agaaacacca	540
	an emperation togethy and an emphasis and endered an emphasis and emph	600
Sagaracâte	cacgtgotte tegettgaag egagttetgg aggatgaaat gaettttaag	3 O U
çaggar aag g	tecetgtatt ggactetaac aaaaggetga aaatgeteea ggateeggtt	660

cgeggagaga	agaaagaagt	aaacgasgga	accasatttg	aatggcttga	gtcttctcga	720
atcagggatg	ccastagaag	acgtectgat	gateceettt	acgatagaaa	gaccttacae	780
ataccacctg	atgicticaa	gaaaatgtct	gcatcacaaa	agcaatattg	gagtgttaag	840
agtgaatata	tggacattgt	gettttett	aasgtgggga	asttttatga	gctgtatgag	900
ctagatgcgg	aattaggtca	caaggagctt	gactggaaga	tgaccatgag	tggtgtggga	96D
asatgcagae	aggttggtat	ctctgaaagt	gggatagatg	aggcagtgca	aaagctatta	1020
gctcgtggat	atamagttgg	acgaatcgag	cagotagaaa	catetgacca	agcaaaagco	1080
agaggtgcta	aractaraat	ccaaggaag	cragiticagg	tättaadled	atcaacagca	1140
адсдадддаа	acategggee	tgatgccqtc	catettettg	ctacasasga	gatcaamatg	1250
gagctacaee	agtgttcaac	tgtgtatgga	tttgcttttg	ttgactgtgc	tgccttgagg	1260
ttttgggttg	ggtccatcag	ngangangna	icatgtgetg	cccttggage	gttattgatg	1320
cadderrepe	caaaggaagt	gttatatzác	agcaaagggc	tatemagaga	agcacaaaag	1330
getetaagga	aatatacgtt	dacsäddrec	acggcggcac	agttggctcc	3gtaccacsa	1440
gtaatggggg	atacagatge	tgctggagtt	agaaatataa	tagaatotaa	cggatacttt	1500
aaaggetett	crgaatcatg	gaactgtgct	gttgatggtc	taaatgaatg	tgztgttgcc	1560
cträgtgete	ttggagaget	asttaatcat	ctgtctagge	taaagctaga	agatgtactt	1620
aagcatgggg	atatttttcc	ataccaaget	tacaggggtt	gcctcagaat	tgatggccag	1580
acgatggtaa	atcttgagat	atttaacaac	agctgtgatg	gtggtectta	agggaccttg	1740
tacaaatate	ttgatametg	tgttagteca	actggtaage	gactottaag	gaattggatc	1800
tgccatccac	tcaaagatgt	agaaagcatc	aataaacggc	ttgatgtagt	tgaagaattc	1860
acggcaaact	cagaaagtat	gcasatcact	ggccagtatc	tocacaaact	tccagactta	1920
gaaagactgc	toggaegeat	caagtotago	gttcgatcat	cagcototgt	gttgcctgct	1980
cttctgggga	aaaagtgct	gaaacaacga	gtraaagcat	ttgggcaaat	tgtgaaaggg	2040
ttcagaagtg	gaattgatct	gttgrrggct	ctacagaagg	aaccaaacat	gatgagtttg	2100
ctctataaac	totgtaaact	tectacatta	gtaggaaaaa	gcgggctaga	gttatttctt	2160
tetcaatteg	AAGCAGCCAL	agatagçg				2195

<210>	27
<212>	1305
<212>	DNA
<213>	Arabidopsis thaliana ecotype Columbia
<223>	Clone 62
c400>	77

catcasecte tgtgttgcet getettetgg ggaaakaagt getgaaacaa cgagttaaag 60 cattigggca aattgigaaa gggitcagaa giggaattga toigiigitg gobotacaga 120 180 aggaatcaaa tatgatgagt ttgctttata aactotgtaa acttcccata tcagtaggaa aaagogggot agagttatti ottiotoaat togaagoago catagatago gactttocaa 240 attatcagaa ccaagatgtg acagatgaaa acgotgaaac totcacaata ottatcgaac 300 360 tttttatega aagageaact caatygtetg aggteattea caccataage egcctagatg toctgagato cootgeaaro geageaagte tototgetgg aageatggee aggeetgtma 42D ttttteeega ateagaaget aeagateaga ateaghaade hooogggeed atdettaaaa 480 tocsaggact acgecatocs titigoagoog cageogotgg teastiffeet giteegaatg 540 600 atatactect tggegagget agaagaagea gtggeageat teatectegg teategttac ់ឥ៩០ tgacgggacc aaacatgggc ggaaaatcaa otettetteg tgcaacatgt ctggccgtta tettigeera actiggetge taegigeegi gigaqielig egaaatetee etegiggata 720 780 ctaccetcae aaggestigge geatebigata gaateacgae aggagagagt accettitigg tagaatgcae tgagacageg teagttette agaatgebae teaggattea etagtaatee 840 900 ttgåcgaact gggcagagga actagtactt tcgatggata cgccattgca tactcggttt 960 CCCGCCACCE ggtagagaaa gttcaatgtc ggatgetett tgcaacacat taccaccete traccangga attracett caccencyty tracctegan acacetyget tycycatten 1020 1080 aatcaagato tgattatoaa comogtggtt gtgatcaaga cotagtgtto ttgtacogtt taaccgaggg agottgtoot gagagotaog gacttoaagt ggcactoatg gotggaatac 1140 caaaccaagt ggttgaaaca gcatcaggtg ctgctcaagc catgaagaga tcaattgggg 1200 1260 addaCttCAA gtcaagtgag Ctaagatctg agttotcaag tctgcatgaa gactggctca agreatiggt gggtatitet cgagtegeec acaacaatge ecceatigge gaagatgact 1320 acqueettt gttttgetta tygeatyaga teaaateete ttactgtgtt cccaaataae 1390

ccggg	1385
<210>	28
c211>	34
<212>	DNA
<213>	Arrificial sequence
	verestoske ardanee
<220>	
<223>	MSH6 specific primer 288 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	26
ateregggtt	atttgggåå¢ acagtaagag gatt 34
<210>	39
<211>	27
<212>	DNA
<213>	Artificial sequence
<2205	
<233>	MSH6 specific primer SB2 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	29
gcgttcgatc	atcagectet gtgttg: 27
<210>	30
<211>	3606
<212>	DNA
<213>	Arabidopsis theliana ecotype Columbia
10437	manatpala malama cootype columnia
<220>	• •
<221>	CDS
<222>	(142) (3468)
<223>	AtMSH6 full-length cDNA and deduced sequence of the encoded polypeptide
<400>	30
aaagttgag	ccetgaggag tategtttee gecattteta egacgeaagg egaaaatttt 60
tggcgccaat	ctttccccc tttcgaatte tctcagctea aaacatcgtt tetctctcac 120
teteteteae	Adttecasaa a atg cag ege cag aga teg att ttg tet tte 171 Met Gln Arg Gln Arg Ser Ile Leu Ser Phe

														99¢ 61y 25		219	
														gtg Val		2 5 7	,
														tcg Ser		315	;
t ct Ser	gtc Val 60	gat Asp	gag Glu	gtt Val	aga Arg	gga Gly 65	acg Thr	gat Asp	act Thr	P.C.S.	eeg Pro 70	gåģ Glu	aag Lya	gtt Val	erg Pro	\$63	ļ
ogt Arg 75	egt Arg	gtc Val	ctg Leu	ccg Pro	tct Ser 80	eja 339	ttt Phe	aag Lys	ecg Pro	gct Ala 85	gaa Glu	tee Ser	gcc Ala	gst Gly	gat Asp 90	411	ļ.
														gtc Val 105		455	3
				Ser										ecg ecg		507	7
aat Asn	yeb	tea Ser 125	Ser	cta Leu	tgt Cys	atg Met	130 Tha	Ala	aat Asn	gat. Asp	ger. Val	att Ile 135	ect Pro	caa Gln	ttt Phe	555	5
		Asn					Gln					Ala		agt Ser	phe Phe	603	3
agt Ser 155	GJY 939	aga Arg	gct	gaa Glu	ctt Leu 160	Àrg	tca Sex	gta Val	gaa Glu	gat Asp 165	Ile	GJA GGS	gta Val	gat Asp	ggc Gly 170	65:	1
					Glu					Arg					Arg	691	9
				Leu					The					Lys	gtt Val	74	7
			Asp					Leu					Asp		gtt Val	79.	5

								аза Lys 230				943
								cgt Arg				891
								gat Asp				939
								aag Lys				987
								tat Tyr				1035
				_			-	119 71p	_	_	_	1083
	-			_	-			tct Şer	_	_		1131
	_				-	_		tat Tyr		_	-	1179
								gcc Ala				1227
								act Thr				1275
								ctt Leu 390			aaa Lye	1323
								gtg Val				1371
								ggg Gly				1419

gac Asp	gcs	tca Ser	tgt Cys 430	gct Ala	got Ala	ctt Leu	G) y gga	gcg Ala 435	tta Leu	ttg Leu	atg Met	cag Gln	gtt Val 440	tct Ser	blo	1467
гув	C]π åss	gtg Val 445	tta Leu	tat Tyr	gaç As p	agt Şer	aaa Lys 450	eja 8aa	cta Leu	tca Ser	aga Arģ	gaā Glu 455	gca Ala	caa Gln	aag Lys	1515
								9JA 8 32								1563
cca Pro 475	gtå Val	cca Pro	caa Gln	gta Val	atg Met 480	Gly ggg	gat Asp	aca Thr	gat Aвр	gct Ala 485	get Ala	GJÀ đàs	gtt Val	aga Arg	aat Asn 490	1611
ata Ile	ata Ile	gaa Glu	tct Ser	aac Asn 495	Gly	сас Туг	ttt Phe	aaa Lys	ggt Gly 500	tct Ser	tet Ser	gaa Glu	tca Ser	tgg Trp 505	aac Asn	1659
tgt Cys	AT9 GCC	gtt Val	gat Asp 510	61y	cta Leu	aut Asn	gaa Glu	S15 Cys	gat Asp	gct Va)	gcc Ala	ctt Leu	agt Ser 520	λla	ctt Leu	1707
gga Gly	gag Glu	cta Leu 525	ıle	aat Asn	cat His	Ctg	tct Ser 530	yrg	cta Leu	aag Lys	cta Leu	gaa Glu 535	gat Asp	gtå Val	Leu	1755
aag Lys	cat His 540	gly	даt	att 11e	ttt Phe	cca Pro 545	îyr Tyr	caa Gln	gtt Val	tac Ţyr	220 yrg 988	Gly	tgt Cys	ctc Leu	aga	1803
att Tle 555	Asp	gjy gge	Gln	aog Thr	atg Met 560	Val	aat	cci Lev	gag Glu	ata Ile 565	5µe	аас Авл	aat Asn	agc Ser	tgt Cys 570	1851
					Gly			tac Tyr		lla					val	1899
agt Ger	cca Pro	act Thr	ggt Gly 590	Lye	. Arg	ctc Led	tta Lev	agg Arg 595	Авл	tgg Trp	r ato	tgc Cya	cat His 600	Pro	ctc	1947
aas Lys	gat Asp	gta Val	. 0ใบ	agc Sei	ato Ile	aat Aen	aaa Lys 610	: Arg	ctt Lev	gat Asp	; gca	gtt Val 619	Glu	gaa Glu	ttc Phe	1995
		Asn					: Ġļī					ו אַנ			aaa Lys	2043

ctt	cca	gac	tta	gaa	aga	ctg	ctc	අපුය	ege	atc	aag	tet	age	9 t E	CGA		2091
											File						
t ca Ser	tca Ser	grc	tct Ser	gtg Val 655	ttg Lev	ect Pro	Ala	ctt Leu	ctg Leu 660	gj ååå	Lys	aaa Lys	gtg Val	ctg Leu 665	aaa Lys		2139
											eja 333						2187
											aat Asn						2235
											gga Gly 71n						2283
											yab gsc						2331
											get Ala						2379
											caa Gln						2427
											tct Ser						2475
											get Val 790						2523
tca Ser 795	Gjn ĝas	gct Ala	aca Thr	gat Asp	Gln Soo	aat Asn	cag Gln	aaa Lys	aca Thr	aaa Lys 805	G1A G33	CCR PYD	ata Ile	ctt Leu	aaa Lys 810	:	2571
atc Ile	caa Gln	gga Gly	cta Leu	tgg Trp 815	cat His	610 CCS	ttt Phe	gca Ala	gtt Val 820	gca Ala	gcc	gat Asp	elá aac	CRA Gln 825	ttg leu		2619
cst Pro	gtt Val	ccg Pro	aat Asn 830	gat Asp	ata Ile	ctc Leu	Cen	gge Gly 835	Gju gag	gct Ala	aga Arg	aga Arg	agc Ser 840	agt Ser	gg¢ Gly		2667

agc Ser	att Lle	cat Ris 845	ect Pro	cgg Arg	tca Ser	r tg Leu	tca Leu 950	ctg Leu	acg	gga Gly	cca Pro	aac Aan 855	atg Met	gg¢ Gly	gga Gly	2715
Lyg	tca Ser 060	art Thr	cct Leu	ctt Leu	egt. Arg	gca Ala 865	aca Thr	tgt Cys	ctg Leu	gcc Ala	gtt Val 870	atc l}e	ttt Phe	gcc	caa Gln	2763
ctt Leu 875	gg¢ Gly	tgc Cys	tac Tyr	gtg Val	ccg Pro 800	tgt Сув	gag Glu	tet Ser	tgc Cys	gaa Glu B85	atc Ile	rcc Ser	ctc Leu	gtg Val	gat Asp 890	3811
act Thr	atc Ile	ttc Phe	aca Thr	agg Arg 895	ctt Leu	G) y gg¢	gca Ala	cct Ser	gat Asp 900	aga Arg	atc Ile	atg Met	aca Thr	gga Gly 905	gag Glu	2859
agt Ser	acc Thr	ctt Phe	ttg Leu 910	gta Val	Ċĵ <i>n</i> Ĝ≅ŝ	tgc Çys	act Thr	gag Glu 915	aca Thr	Ala	tca Şer	gtt Val	ctt Leu 920	cag Gln	aat Asn	2907
gca Nla	act Thr	cag Gln 925	gat Asp	tca Ser	cta Leu	gta Val	atc Ile 930	ctt Leu	gac Asp	gaa Glu	i.ev ccg	935 935	yrd	Gly	act Thr	2955
agt Şer	act Thr 940	Phe	gat Asp	gga Gly	cac Tyr	gcc Ala 945	acc Ile	gca Ala	tac Tyr	t eg Ser	get Val 950	БуБ	ogt Arg	cac His	ctg Leu	3003
gta Val 955	GJ <i>n</i> 393	aaa Lys	gtt Val	caa Gln	tgt Cys 960	Arg	atg Met	at.c Leu	ttt Phe	965	The.	cat His	tac Tyr	сас Нів	970	3051
cto Leu	acc Thr	aag Lye	gaa Glu	ttc Phe 975	gcg Ala	t¢t Ser	cac His	cca Pro	ogc Arg	Væl	acc Thr	teg Ser	asa Lye	cac His 985	atg : Met	3099
gct Ala	tgc Cys	gca Ala	ttc Phe 390	Lys	tca Ser	sgs gr4	tct Ser	gat Asp 995	тут	caa Gln	e ecu	cgt Arg	: ggt : Cly : 1000	Сув	gat Asp	3147
Cln	gac	cta Leu 1005	val	ttc Phe	ctg Lev	tac Tyr	cgt Arg	Leu	acc Thr	gag Glu	. Gly 1 gga	gct Ala	Cys	cct Pro	gag Olu	3195
ser	tac Tyr 1020	Gly	ctt Leu	caa Glm	gtg Val	gea Ala 1025	Lev	atg Met	gct Ala	gga Gly	ata / Ile 1030	Pro	aac Asr	cae 1 Oli	a gtg 1 Val	3243
gtt Val 103	Glı	acs Thr	gca Ala	tea Ser	ggt : Gly	/ Ala	get Ala	caa Gla	gco Ala	e atq 1	Lys	a yrd 1 sg:	a tea g Sei	att CIle	ggg Gly 1050	3291

gga Glu	aac Asn	tto Phe	Гላጻ	tca Ser 1055	agt Ser	GJ <i>n</i> Bait	cta Leu	Arg	tet Ser 1060	gag Glu	ttc Phe	tca Ser	ser	ctg Leu 1065	oat Xis	3339
gaa Glu	gac Asp	ı, tb	ctc Leu 1070	аад Гув	cca Ser	ttg Lev	Val	ggt Gly 1075	att Ile	tct Ser	cga Arg	Val	gee Ala 1080	cac Xis	aac Asn	3387
aat Aen	Ala	CCC Pro 1085	atț Ile	gjy âāc	Gj <i>n</i> Gaa	Узb	7090 Asb Ggc	tac Tyr	yab dac	act Thr	Leu	ttt Phe 1095	tgs	tta Lev	1rp 193	3435
Ris	gag Glu 1100	atr	rys	rcc	Ser	tac Tyr 1105	tgt Cys	gtt Val	ccc Pro	aaa Lys	taai	atgg	cta			3478
tga	cata.	aca (ctate	ctgaa	ag c	tcgt1	taagi	t cti	ttg	cctc	tet	gatg	ttt i	arte	ctccta	3538
388	aatg	ett d	stata	et caa	કાર શક	aikt-Li	ett:	. ಆರಂ	gac	caaa	યસથ	2230	aaa l	3388	Baaaaa	3598
282	BBBB	D				,										3606
<210 <211 <311 <211 <221 <401	1> 2> 3>			r zbiđa		s the ≥ MS}		na ek	cotyr	රාළ ඉර	> 1 ນ ຕ ນ້	oia				
Met 1	Gln	Arg	Gln	Arg 5	Ser	île	гел	Ser	Dps 70	Phe	Gln	Lув	Pro	Thr 15	Ala	
Ala	Thr	Thr	Lys 20	Gly	ren	Val	Ser	Gly 25	АБĎ	Ala	Ala	Ser	Gly 30	Gly	Gly	
Gly	Ser	Gly 35	GJY	Pro	Arg	Phe	A\$D 40	val	Arg	CJA	gly	Лар 45	Ala	Lys	Gly	
Asp	Ala 50	Ser	Va1	Arg	Phe	Ala 55	val	ser	Lys	Ser	Val 60	Asp	Glu	Val	yr.ü	
G1y 65	Thr	двд	Thr	Pro	Pro 70	Glu	Lys	Val-	Pro	Arg 75	Arg	Va1	Leu	Pro	9er 80	
Gly	Phe	Lys	Pro	Ala 85	Glu	ser	Ala	Ġĵλ	Авр 90	Ala	Ser	Ser	Leu	Phe 95	Ser	
Asn	lle	Met	eiR 100	Lya	Phe	Val	Lya	Val 105	Asp	Азр	Arg	Asp	Cys 110	Ser	Gly	,
СJЛ	Arg	Ser 115	yrg	дји	Asp	Val	Val 120	Pro	Leu	asn	qeA	Ser 125	ser	Leu	Суз	

Met Lys Ala Asn Asp Val Ile Pro Gln Phe Arg Ser Asn Asn Gly Lys 135 130 Thr Sin Glu Arg Ash His Ala Phe Ser Phe Ser Gly Arg Ala Glu Leu 150 195 145 Arg Ser val Glu Asp Ile Gly Val Asp Gly Asp Val Pro Gly Pro Glu 17D Thr Pro Gly Met Arg Pro Arg Ala Ser Arg Leu Lys Arg Val Leu Glu 185 Asp Glu Met Thr Phe Lys Glu Asp Lys Val Pro Val Leu Asp Ser Asn 195 Lys Arg Leu Lys Met Leu Gla Asp Pro Val Cys Gly Glu Lys Lys Glu 215 Wal Asn Glu Gly Thr Lys Phe Giu Trp Lon Clu Sor Ser Arg Ile Arg 230 235 Asp Ala Ash arg Arg Arg Pro Asp Asp Pro Leu Tyr Asp Arg Lys Thr 250 255 245 Lou Bis The Pro Pro Asp Val Phe Lys Lys Met Ser Ala Sor Glo Lys 265 260 Glp Tyr Trp Ser Val Lys Ser Glu Tyr Met Asp Ile Val Leu Phe Phe Lys Val Gly Lys Phe Tyr Glu Leu Tyr Glu Leu Asp Ala Glu Leu Gly His Lys Glu Lou Asp Trp Lys Met Thr Met Sor Gly Val Gly Lys Cys 310 315 305 Arg Gln Val Gly Ile Ser Glu Ser Gly Ile Asp Glu Ala Val Gln Lys 325 33D Leu Leu Ala Arg Gly Tyr Lys Val Gly Arg Ile Glu Gln Leu Glu Thr Ser Asp Cin Ala Lys Ala Arg Cly Ala Asn Thr Ile Ile Pro Arg Lys 355 Leu Val Glm Val Leu Thr Pro Ser Thr Ala Ser Glu Gly Asn Ile Gly 3 B Q 370 375 Pro Asp Ala Val His Leu Leu Ala Ile Lys Glu Ile Lys Met Glu Leu 395 390

Gln Lys Cys Ser Thr Val Tyr Gly Phe Ala Phe Val Asp Cys Ala Ala

- Leu Arg Phe Trp val Gly Ser Ile Ser Amp Asp Ala Ser Cys Ala Ala 420 425 430
- Leu Gly Ala Leu Leu Met Gln Val Ser Pro Lys Glu Val Leu Tyr Asp 435 440 445
- Ser Lys Gly Leu Ser Arg Glu Ala Gln Lys Ala Leu Arg Lys Tyr Thr 450 455 460
- Leu Thr Gly Ser Thr Ala Val Gin Leu Ala Pro Val Pro Gin Val Met 465 470 475 480
- Gly Asp Thr Asp Ala Ala Gly Val Arg Asn Ile Ile Glu Ser Asn Gly
 485 490 495
- Tyr Phe Lys Gly Ser Ser Glu Ser Trp Asn Cys Ala Val Asp Gly Leu 500 505 510
- Asn Glu Cys Asp Val Ala Lou Ser Ala Leu Cly Glu Leu Ele Asn His 515 520 525
- Leu Ser Arg Leo Lys Leo Giu Asp Val Leo Lys His Gly Asp The Phe 530 535 540
- Pro Tyr Gln Val Tyr Arg Gly Cys Leu Arg Ile Asp Gly Gln Thr Met 545 550 555 560
- Val Asn Leu Glu Ile Phe Asn Asn Ser Cys Asp Gly Gly Pro Ser Cly 565 570 575
- The Leu Tyr Lys Tyr Leu Asp Asn Cys Val Ser Pro The Gly Lys Arg
- Leu Leu Arg Asn Tro Ile Cys Ris Pro Leu Lys Asp Val Glu Ser Ile 595 600 605
- Asn Lys Arg Leu Asp Val Val Glu Glu Phe Thr Ala Asn Ser Glu Ser 610 620
- Met Gln Ile Thr Gly Gln Tyr Leu His Lys Leu Pro Asp Leu Glu Arg 625 630 635 640
- Leu Leu Gly Arg Ile Lys Ser Ger Val Arg Ser Ser Ala Ser Val Leu 645 650 655
- Pro Ala Leu Leu Gly Lys Lys Val Leu Lys Gln Arg Val Lys Ala Phe 660 665 670
- Gly Gln Ile Val Lys Gly Phe Arg Ser Gly Il Asp Leu Leu Ala 675 680 585
- Leu Gln Lys Glu Ser Aan Met Net Ser Leu Leu Tyr Lys Leu Cys Lys 690 595 700

STOCIF-WO GOLDADATTI

- Leu Pro Ile Leu Val Gly Lys Ser Gly Leu Glu Leu Phe Leu Ser Gln 705 710 715 720
- the Glu Ala Ala Ile Asp Ser Asp Phe Pro Asm Tyr Cln Asm Glm Asp 725 730 735
- Val Thr Asp Glu Asn Ala Glu Thr Lev Thr He Leu He Glu Leu Phe 740 745 750
- Ile Glu Arg Ala Thr Gln Trp Ser Glu Val Ile His Thr Ile Ser Cys 755 760 765
- Leu Asp Val Leu Arg Ser Phe Ala Ile Ala Ala Ser Leu Ser Ala Gly
 770 775 780
- Ser Met Ala Arg Pro Val Ile Phe Pro Glu Ser Glu Ala Thr Asp Glu 785 790 795 800
- Ann Glm Lys Thr Lys Gly Pro Tie Leu Lys Ile Glm Gly Leu Trp Mis 805 810 815
- Pro Pho Ala Val Ala Ala Asp Ely Glm Leu Pro Val Pro Asn Asp Ile 820 825 830
- Let Let Gly Glu Ala Arg Arg Ser Ser Gly Ser Ile His Pro Arg Ser 835 840 845
- Leu Leu teu Thr Gly Pro Asm Met Gly Cly Lys Ser Thr Leu Leu Arg
- Ala Thr Cys Leu Ala Val Ile Phe Ala Gln Leu Gly Cys Tyr Val Pro 865 870 875 880
- Cys Glu Ser Cys Glu Ile Ser Leu Val Asp Thr Ile Phe Thr Arg Leu 885 890 895
- Gly Ala Ser Asp Arg Ile Met Thr Gly Glu Ser Thr Phe Leu Yal Glu 905 910
- Cys Thr Glu Thr Ala Ser Val Leu Gln Asn Ala Thr Gln Asp Ser Leu 915 920 925
- Val Ile Leu Asp Glu Leu Gly Arg Gly Thr Ser Thr Phe Asp Gly Tyr 930 935 940
- Ala Ile Ala Tyr Ser Val Phe Arg His Leu Val Glu Lys Val Gln Cys 945 950 955 960
- Arg Met Leu Phe Ala Thr His Tyr His Pro Leu Thr Lys Glu Phe Ala 965 970 975
- Ser His Pro Arg Val Thr Ser Lys His Met Ala Cys Ala Phe Lys Ser 980 985 990

Arg Ser Asp Tyr Gln Pro Arg Gly Cys Asp Gln Asp Leu Val Phe Leu 995 1000 1005	
Tyr Arg Leu Thr Glu Gly Ala Cys Pro Glu Sex Tyr Gly Leu Gln Val 1010 1015 1020	
Ala Leu Met Ala Gly Ile Pro Asn Gln Val Val Glu Thr Ala Ser Gly 1025 1030 1035 1040	
Ala Ala Gin Ala Met Lys Arg Ser Ile Gly Glu Asn Phe Lys Ser \$er 1045 1050 1055	
Glu Leu Arg Ser Glo Phe Ser Ser Leu His Glu Asp Trp Leu Lys Ser 1060 1065 1070	
Leu Val Gly Ile Ser Arg Val Ala His Asn Amn Ala Pro Ile Gly Glu 1075 1080 1085	
Asp Asp Tyr Asp Thr Leu Phe Cys Leu Typ His Glu Ile Lys Ser Ser 1090 1095 1100	
Tyr Cys Val Pro Lys 1105	
<210> 32	
<211> 24	
<212> CNA	
<pre><213> Artificial sequence</pre>	
<220>	
<223> Forward primer for PCR amplification of ATRGENEA	
microsatellite	
<400> 32	
accatgoata gottaaactt ottg	24
<210> 33	
<211> 22	
<212> DNA	
<213> Artificial seguence	
<220>	
<223> Reverse primer for PCR amplification of ATHGENEA microsatellite	
<400> 33	
acataaccac aaataggggt oc	22

<210>	34			
<211>	18			
4212>	ona			
<213×	Artificial sequence			
<220>				
<223>	Porward primer DMCIN-A for PCR on go	enomic :	DNA 4	of Arabidocsis
	thaliana ssp. Landsberg erecta *Ler			
<400>	34	•		
gaagogatat	tqttcqtq		•	18
<510>	35			
<211>	13			
<212>	DNA			
c213>	Artificial sequence			
<220>				
<223>	Reverse primer DMCIN-B for PCR on g	enomic	DNA (of Arabidopaia
	thaliana ssp. Landsberg erecta "Ler			
<400>	35	•		
agattqcgag	aacattct			18
<210>	36			
<211>	31			
<212>	DNA			
<213>	Artificial sequence			
<220>				
<223>	Forward primer DMCIN-1 for PCR on g thaliana ssp. Landaberg erecta "Ler		DNA	of Arabidopsis
<400>	36		•	
acgcgtcgac	tcagctatga gattactcgt q			31.
<21 0 >	37			
<211>	29			
<212>	DNA			
<213>	···			
488J2	Artificial sequence	_		
<220>				
c273>	Reverse primer DMCIN-2 for PCR on g thaliana ssp. Landsberg erecta "Ler		DNA	of Arabidopsis
<400>	37			
getetagate	tetegeteta agactetet			29

<510 >	38
<211>	32
<212>	DNA
~213×	Artificial sequence
<220>	
<233>	Forward primer DMCIN-3 for PCR on genomic DNA of Arabidopsia thaliana sap. Landsberg erecta "Ler"
<400>	38
g¢tstagag	t ttctcttaag taagtgattg at 32
<210>	39
<211>	49
<212>	DNA
c2135	Artificial seguence
<2205	•
<223>	Reverse primer DMCIN-4 for PCR on genomic DNA of Arabidopsis chaliana sep. Landsberg erecta "Ler"
<400>	39
receerggy)	> togaçayato tocatggiti oficagetor algastoc 48
<210>	40
<2115	26
<2125	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer DMC1a for PCR on genomic DNA of Arabidopsis theliana ssp. Landsberg erects "Ler"
<400>	40
acycytegad	gaattegeaa gtgggg 26
<210>	41
<211>	38
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer DMC1b for PCR on genomic DNA of Arabidopsis thalians sap. Landsberg erecta "Ler"
-400-	45

FCCatggaga	totecogggt accgatetge ttcgaggg 38
<210>	42
<212>	20
<212>	DNA
<213>	Artificial sequence
-6157	
<220>	
<223>	Forward primer for PCR amplification of ATBAT1 SSLP marker in Arabidopsis thaliana subapecies
<400>	42
gccactgcgt	gastgatatg 20
<210>	4.3
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of ATEAT1 SSLP marker in Arabidopsis thaliana subspecies
<400>	43
ogaacageca	scattaattc cc 22
<210>	44
<211>	18
<212>	DNA
<213>	Artificial sequence
70207	ntite to the conduction
<220>	
<223>	Forward primer for PCR amplification of NGA63 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	44
aaccaaggca	cagaagcg
<210>	45
<211>	18
<211>	DNA
<213>	Artificial sequence
-01J/	NA CASSCARIA ROGUESTICO
<220>	
<223>	Reverse primer for PCR amplification of NGA63 SSLP marker in Arabidopsis thalians subspecies

<400>	45	
acccaagtga	Logocace .	18
		(
<210>	46	
<211>	21	
<213>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of NGA248 SSLP marke	r in
	Arabidopsis thaliana subspecies	
<400>	46	
taccgaacca	asacacsaag g	21
<210>	47	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
	·	
<220>		
<223>	Reverse primer for PCR amplification of NGA248 SSLP marke Arabidopsia thaliana subspecies	r in
<400>	47	
tototatoto	ggtgaattet cc	22
	••	
<210>	48	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
<220>	·	
<223>	Forward primer for PCR amplification of NGA128 SSLP marks	r in
	Arabidopsis thalisma subspecies	
<400>	48	
ggtctgttga	Egtegraagt cg	22
<210>	49	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
-	·	
<220>		

<223>	Reverse primer for PCR amplification of NGA128 SSLP marker in Arabidopsis thalians subspecies
<400>	49
atottgaaac o	tttagggag gg 22
·	
<210>	50
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA280 SSLP marker in Arabidopsis thaliana subspecies
<400>	50
ctgatctcac g	gacaatagt gc 22
<210>	51
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA280 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	51
arakeantoo -	120 aaagcgcacc
ggctccataa a	raagegenee 20
<210>	52 .
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGAll1 SSLP marker in
	Arabidopsis thaliana subspecies
<\$00>	52
	31
orneadring s	aagctaaagg g 21
	•
c210>	53
<211>	21
<212>	DNA
<213>	Artificial sequence

<220>		
<223>	Reverse primer for PCR amplification of NGAlll SSLP marker Arabidopsis thaliana subspecies	in
<40D>	53	
tgttttttag	gacaaatggc g 2	11
<210>	54	
<211>	20	
<212>	DNA	
c213>	Artificial sequence	
<220>	*	
c223 >	Porward primer for PCR amplification of NGA168 SSLP marker Arabidopsis thaliana subspecies	iŋ
<400>	54	
cetteacate	саадассевс 2	20
<210>	55	
5211>	20	
<212>	DNA	
<213>	Artificial sequence	
<220>		
c223>	Reverse primer for PCR amplification of NGA168 SSLP marker Arabidopais thalians subspecies	in
<40D>	55	
gcatataccc	acaaccagaa	20
<210>	56	
<211>	20	
<212>	DNA	
<213>	Artificíal sequence	
<220>		
<223>	Forward primer for PCR amplification of NGA1126 SSLP marker in Arabidopsis thaliana subspecies	<i>:</i>
<400>	56	
cgctacgctt	ttoggtasag	20

<210>	57
<211>	20
<212>	אמט
<213b	Artificial sequence
	• '
<220>	
€223>	Reverse primer for PCR amplification of NGA1126 SSLP marker
	in Arabidopsis chalians subspecies
	1.5-4-4-4-6-6-0 OWESTERN AMORAGES
<400>	57
11000	
goacagtoca a	gtercaeco 20
20000490000	70
,	
<210>	58
<211>	20
<212>	DNA
<213>	Artificial sequence
<8T3>	Artificisi seddeuce
<2 20 >	
	there are a second and a second and a second are a second as a
<223>	Forward primer for PCR amplification of NGA361 SSLP marker in
	Arabidopsis Chuliana subspecies
4 8 6	
<400>	58
aaagagatga g	aatttggsc 20
*40	M A
<210≻	59
<211>	23
<212>	DNA
<213>	Artificial sequence
<250>	
<223>	Reverse primer for PCR amplification of NGA361 SSUP marker in
	Arabidopsis thaliana subspecies
	ş
<400>	59
acatatcaar a	tattaaagt agc 23
	·
<210>	60
<211>	18
<212>	DNA
<213>	Artificial sequence
	•
<220>	
<223>	Forward primer for PCR amplification of NGA168 SSLP marker in
	Arabidopsis thaliana subspeci s
	•
<400>	60

ccatcractd	cactgecg 18
244	
<210×	61
<211>	ZZ DNA
<212><213>	
<5172	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA168 SSLP marker in Arabidopsis thaliana subspecies
<400>	61
gaggacatgt	ataggageet eg
<210>	-
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Porward primer for PCR amplification of AthBID2 SSL2 marker
	in Arabidopsis Thaliana subspecies
<400>	62
tgacctcccc	ttccatggag 20
<210>	63
<211>	22
<212>	DNA
<213>	Artificial sequence
<320>	
<223>	Reverse primer for PCR amplification of AthBIO2 SSLP marker
	in Arabidopsis thaliana subspecies
<400>	63
ttaacagaaa	cccaaagett te 22
<210>	64
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Porward primer for PCR amplification of AthURIQUE SSLP marker
	in Arabidopsis thaliana subspecies

, 44QDp	64
aggcaaatgt	coattteatt g
<2105	65
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223	Reverse primer for PCR amplification of AthUBIQUE SSLP marker in Arabidopsis thaliana subspecies
<400>	65
acgacatggc	agatttctcc 20
<210>	66
<211>	21
<212>	DNA
<213>	Artificial sequence
<22D>	
<223>	Forward primer for PCR amplification of NGA172 SSLP marker in Arabidopsis thaliana subspecies
<400>	66
ag ctgc ttcc	ttatagegte e 21
<210>	67
<211>	19
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA172 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	67
catorgaatg	ccattgtc 19
<210>	68
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	

```
< 323×
              Forward primer for PCR amplification of NGA126 SSLP marker in
              Arabidopsis thaliana subspecies
<400 >
              58
gasamasege tactttegtg g
                                                                           21
<210>
             69
<211>
             22
<212>
             DNA
<213>
             Artificial seguence
<220×
             Reverse primer for PCR amplification of NGA126 SSLP marker in
<223>
             Arabidopsis thaliana subspecies
<400>
             ಕ 9
caagageaat atcaagagea go
                                                                           22
<210×
             70
<211>
             23
<212>
             DMV
<213>
             Artificial sequence
<220>
<223>
             Forward primer for PCR amplification of NGA162 SSLP marker in
             Arabidopsis thaliana subspecies
<400>
             70
catgomattt geatergagg
                                                                           20
<210>
             71
<211>
             22
<212>
             DNA
<213>
             Artificial sequence
<220>
<223>
             Reverse primer for PCR amplification of NGA162 SSLP marker in
             Arabidopsis thaliana subspecies
<400>
             71
ctctgtcact cttttcctct gg
                                                                           2.2
<210>
             72
<211>
             21
<212>
             ANG
<213>
             Artificial sequence
```

<220>	•	
<333>	Forward primer for PCR amplification of NGA6 SSLP marker i Arabidopsis thaliana subspecies	n
<400>	72	
tggatttett c	ctetettea e	21
<210>	73	
<311>	21	
<212>	DNA	
<313 >	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA6 SSLP marker i Arabidopsis thaliana subspecies	n
<400>	73	
atggagaagc t	tacactgat c	31
<210×	74	
<211>	20	
<212>	DNA	
<213>	Artificial sequence	
<220×		
<2235	Forward primer for PCR amplification of NGA12 SSLP marker Arabidopsis thaliana subspecies	in
<400>	74	
aatgttgtee t	cccctcctc	20
<210>	75	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA12 SSLP marker	in
	Arabidopsis thaliana subspecies	
<400>	75	
tgatgetete t	gaaacaaga gc	22

<210>	76
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCP applification of MGRA GOLD
	Forward primer for PCR amplification of NGA8 SSLP marker in Arabidopsis thaliana subspecies
<400>	76
gaggcasat c	tttattteg g
<210>	77
<211>	22
<212>	DNA
<213>	Artificial sequence
	in this seddelife
<220>	
<323>	Payeren number for DOD
	Reverse primer for PCR amplification of NGA8 SSLP marker in Arabidopsis chaliand subspecies
<400>	77
tggctttegt t	tatasacat cc 22
<210>	78
<211>	21
<212>	DNA
<213>	
4417	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA1107 SSLP marker
	in Arabidopsis thaliana subspecies
<400>	78
gogamasasac a	assastcc a 21
•	
<210>	79
<211>	21
<212>	DNA
<213>	
~ 46 16 17 17	Artificial sequence
<220>	
<223>	Reverse primer for DCP applification of process
	Reverse primer for PCR amplification of NGA1107 SSLP marker in Arabidopsis thaliana subspecies
<400>	79

cdatâ#sccd	Acagaattag g	21
<210>	8 C	
<211>	21	
<212>	DNA	
c213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of NGA225 SSLP marks Arabidopsis thaliana subspecies	er in
<400>	80	,
gaaatccaaa	toccagagag g	21
<210>	B1	
<211>	22	
<212>	DNA	
<213>	Artificial Bequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA225 SSLP marks Arabidopsis thaliana subspecies	∍r in
<400>	81	
totocccact	agttttgtgt cc	22
<210>	B2	
<311>	19	
c212>	ONA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of NGA249 SSLP marke Arabidopsis thaliana subspecies	er in
<400>	82	
taccgtcaat	ttcatcgcc	19
c210>		•
<211>	83 22	
<212>	DNA	
<213>		
7017	Artificial sequence	
<220>		
<2233	Reverse primer for PCR amplification of NGA249 SSLP marke Arabidopsis thaliana subspecies	r in
	• • • • • • • • • • • • • • • • • • • •	

```
<400%
              83
ggatecetas etgtasaate eg
                                                                           22
<210>
              84
<211>
              22
<212>
              DNA
<213>
              Artificial sequence
<220>
<223>
              Forward primer for PCR amplification of CA72 SSLP marker in
              Arabidopsis thaliana subspecies
<400>
eateccagta accasacaca ca
                                                                           22
<210×
              85
<211>
              20
<212>
             DNA
د2135
             Armificial Requence
<220>
<223>
             Reverse primer for PCR amplification of CA72 SSEP marker in
             Arabidopsis challanz subspecies
<4005
             85
cccagtitaa ccacgaccac
                                                                           20
<210>
             86
<2115
             20
<212>
             DNA
<213>
             Artificial sequence
<220>
<223>
             Forward primer for PCR amplification of NGA151 SSLP marker in
             Arabidopsis thaliana subspecies
<400>
             86
gtttcgggaa gttttgctgg
                                                                           20
<210>
             87
<211>
             24
<212>
             DNA
<213>
             Artificial sequence
<220>
```

<223×	Reverse primer for PCR amplification of NGA151 SSLP marker in Arabidopsis thalians subspecies
< 400 >	87
cagtetaaaa	gogagagtat gaig 24
<310>	88
<211>	27
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Porward primer for PCR amplification of NGA106 SSLP marker in Arabidopsis thaliana subspecies
<400>	98
gttatggagt	ttctagggca cg 22
<210×	. 28
<211>	20
<212>	AND
<213>	Artificial sequence
(A 4 3 >	vicilions seddence
<230>	
<223×	Reverse primer for PCR amplification of NGA106 SSLP marker in Arabidopsis thaliana subspecies
<400>	B 9
tgccccarre	tgttcttctc 20
<210>	90 .
<311>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA139 SSLP marker in
	Arabidopais thalians subspecies
<400>	90
agagocacca	gatccgatgg 20
<210>	91
<211>	21
<2115	DNA
<213×	Artificial Sequence
~~ ~~ ~	ur nassassa sedatitt

<220×				
<223>	Reverse primer for PCR amplification of Arabidopais thaliana subspecies	of NGA139	sslp mar	ker in
<400>	91			
ggtttegttt	cactatecag g			21
<210>	93			
<211>	22			
<212>	DNA			
<213>	Artifícial sequençe			
<220>		•		
<223>	Forward primer for PCR amplification of Arabidopsis thaliana subspecies	f NGA76	SSLP mark	ni re
<400>	92			
ggagaaatg	tcactctcca cc			22
<210>	93			
<211>	20			
<212>	DNA			
<213>	Artificial sequence			
<230>				
<223>	Reverse primer for PCR amplification of Arabidopais thaliana subspecies	f NGA76	SSLP mark	er in
<400>	93			
aggcatggga	gacatttacg			20
. 27 0 -	2.0			
<210><211>	94 20			
<211>	DNA			
<213>	•			
(4433	Artifícial sequence			
<220>				
<223>	Forward primer for PCR amplification o in Arabidopsis thaliana subspecies	f ATHSO1	91 65 LP TW	arker
<400>	94			
ctccaccaat	catgesastg			20

	• •
<210>	95
<211>	21
c212>	DNA
<213>	Artificial sequence
c220>	•
<233>	Reverse primer for PCR amplification of ATHSO191 SSLP marker
	in Arabidopsis thaliana subspecies
<400>	95
tgatgttgat	ggagatggtc a 21
	-
<2105	96
c211>	22
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA129 SSLP marker in Arabidopsis thatiana subspecies
<400>	96
tcaggaggaa	ctatagtgag gg 22
<210>	97
<211>	22
<212>	DNA .
<213>	Artificial sequence
<2 20 >	
<223>	Reverse primer for PCR amplification of NGA129 SSLP marker i Arabidopsis thaliana subspecies
<400>	97
cacactgaag	atggtettga gg 22
<210>	98
<211>	8062
<212>	DNA
<213>	Arabidopsis thaliana ecotype Columbia
<220>	
<223>	Ganomic DNA sequence of AtMSR6
<400>	97

tittitggtt gct acaata asggtatace gttttatgtc atcastataa ctatatataa 60

aagaaasgaa	ayatatatat	tgtttttcca	tttatcaaac	86364868	gastttttt	120
ttacttttta	cattggteaa	casaatacaa	gatasacgac	accetttaat	catttcccaa	190
ttttacccet	aagtttaaca	cctagaacct	tetecatett	egcaageaea	geetgattag	240
gaəcagcttt	accattetca	tatteetgaa	ctacctgagt	cctctrattg	atctgtttcg	300
ccaaatccgc	ttgtgacatc	ttottotoca	atctcgctct	cogtatoato	aacctcacct	360
ctgctttcac	acgatecate	gccgcaggrt	etgittette	ttccagette	ttcgtgttaa	420
ttacoggaac	cgccgtagat	ttcccctttt	tgttcgaacc	ggcatcgaat	ttettaaccg	480
tttgaaccgc	gacacegttt	ctcagagetg	egttaacege	ttteggateg	cgtaggtett	54D
ggctcttttg	ttttgatttg	tggagaacta	ctggttccca	gtettgtgtt	actgctcctg	E00
ggtatctgct	cggcatcqtc	gatgasttga	gagaaagqaa	саасдстава	ättttattaa	660
tetgagettt	gaaattgaga	aacgatgaaq	atgaagaatg	ttgttgagag	gattqtgata	720
tttateceta	cgaagatçgg	tttctggaga	attegateac	crettectec	attttcgtct	780
ctggaacgtt	cttagagatg	attgacgacg	tgtcattate	rgatttgcag	ttaaccaatg	840
ctttttgggt	tggattcgtg	gtacaccata	ttatccgatt	tggctcaatg	gttttätata	900
aatteggete	toggtteggt	tatgagttat	Cattaeaatt	aagetaacea	aaaattttcg	960
taaaatttat	tteggtttca	attoggator	ettaetteea	gaaccgaatt	attegaaace	1020
ggggcragee	gaaccgaata	ccsatgcccg	attgactcgc	tggotagaaa	gatccaacgg	1080
tatecaataa	tagaacatas	ateggaeggt	cateaaagee	tcaaagagtg	sacagtcasc	1140
amamamagtt	gagccctgag	gagtatogtt	teegecatet	ctacgacgea	aggcgaaaat	1200
ttttggcgcc	aatettteec	ccctttcgaa	tteteteage	tcaaaacatc	gtttetetet	1260
CBStStStStSt	cacaattcca	aaaaatgcag	cgccagagat	cgattttgtc	tttcctccaa	1320
asacccacgg	cggcgactac	gmagggtttg	gtttccggcg	atgetgetag	çG GÇGG GGÇ	1380
adcødcådød	accacgattt	aatgtgaagg	aaggggatgc	caaaggcgac	gettetgtae	1440
gttttgctgt	ttcgaaatct	gtcgatgagg	ttagaggaac	ggatactoca	ccddedaagg	1500
ttccgcgtcg	tgtcctgccg	tctggattta	agccggctga	atccgccggt	gatgettegt	1560
ccctgttctc	caatattatg	cataagtttg	taaaagtega	tgatogagat	tgttctggag	1620
agaggtecta	atottogatt	ctcttaattt	tgttatettt	agetggaaga	agaagatteg	7680

tgtaatttgt	tgtattcgtt	ggagagatte	tgattactgc	sttggategt	rgretacasa	1740
tittcaggag	cogagaagat	gttgttccgc	tgaatgattc	atototatgt	atgaaggeta	1800
atgatgttat	teetraattt	cgttccaaca	atggtaaaac	tcaagaaaga	aaccatgctt	1860
ttagrttcag	tgggagaget	gaacttagat	cagtagaaga	tateggagta	gatggcgatg	1920
ctcctggtcc	agaaacacca	dada racarc	cacgtgotte	tegettgaag	cgagttctgg	1980
aggatgaaat	gactittaag	gaggataagg	ttretgtatt	ggactetaac	aaaaggctga	2040
aaatgctcca	ggatccggtt	tgcggagaga	agaazgaagt	ssacgasggs	accaaatttg	2100
aatggettga	gtettetega	atcagggatg	ccsatagaag	Acgrectigat	gatcecettt	2160
acgatagaaa	gaccttacac	ataccacctg	atgttttcaa	gamaatgtot	gcatcacasa	2220
egcaatattg	gagtgttaag	agtgaatata	tggacattgt	getttettt	aaagtggtta	2280
gtaactatta	atctagtgtt	caatccatct	cctcaatgcg	atttgttcac	ttacatotgt	2340
ttacgttatg	CtCttCtCag	gggaaatttt	atgagotgta	tgagctagat	gcggaattag	2400
gtcacaagga	gcttgactgg	sagacgaeca	tgagtggtgt	gggaaaatgc	agacaggtaa	2460
attagttgaa	acaactggcc	tgottgaatt	attgtgtcta	taaattttga	caccaccttt	2520
tgtttcaggt	tggtatetet	d sse &£ddds	tagatgaggc	agtgcaaaag	ctattagete	2580
araaaraada	gaaccatcat	actttatgga	attegtttac	tgctacttcg	gctaggattt	2640
a aga aatgga	aatcacttca	agcatcatta	gttaggatcc	tgagaactca	ggatgttttc	2700
ttättegtta	tataataagt	cttttcatca	#ddadtests	aacaaaactt	gcacaatatt	2760
tgtgtgctca	ctggcaaggr	atatatacce	agctaacctt	tgctagttca	ctgtagtaac	2820
agttacggat	&&E&Eatgtt	tactrgtatg	tggtaccetc	attitgtete	tcatggaggc	2880
tttcaagcct	tgtgttgasa	ctggacagtt	acatátýctt	ссавсадаав	ctagcatgca	2940
gattcatatg	ctttcctatt	ctactaatta	tgtattgaca	cactcgttgt	ttcttttgaa	3000
agatataaag	ttggacgaat	egagcageta	gaaacatetg	ассаадсааа	agceagaggt	3060
gctaatactg	taagttttct	tggataggtc	aaggagagtg	ttgcagactg	tttttgatca	3120
tttcttttc	tgtacattac	tttcatgctg	taattaactc	aatggctatt	ctggtctgat	3180
tatcagataa	CCCCaaggaa	gctagttcag	gtattaacte	catcaacagc	aagegaggga	3240
aacatogggc	ctgatgecet	ccatcttctt	gctataasag	Agenthelita	tttacttatt	3300

tatettatea	tgttcagttc	Atccangtcc	tgassastta	cactottott	taecastett	3360
ccatcaaget	gtytäääggä	tttggaatta	gaaaatcatt	atttgatgct	ttgttttata	3420
tgcaagaggt	tcccttgaaa	agacctgttt	aagattettt	gcacttgaaa	anticaatct	3480
ttttaagtga	atcrectact	ttcttacaat	gatcatagtc	tgcsattgca	tgtcmagtma	3540
tatcattect	tgttactgca	tecepetett	tcttaatgae	cattgtctat	gttgtgtttg	3600
tetegtgtge	tggagaaaat	gatagetgat	ccaagctgta	cattatcatg	attaagtagr	3660
tgct caggaa	ttgcctttgg	ttacattgee	taatggtttg	atgtcaattt	ttcttctgaa	3720
tctttattt	agatcaasat	ggagctacaa	aagtgttcaa	ctgtgtatgg	atttgctttc	3780
gttgaetgtg	ctgccttgag	gttttgggtt	gggtccatca	gcgatgatgc	atcatgtgct	3840
getettggag	cgttattgat	gcaggtaagc	aagtgtattc	tgtatcttat	gtgtaccatg	3900
tgacttcctg	tgcatatatt	tgggttgcag	gaactaattc	tgaatcacca	tttggtatgt	3960
tttttccagg	tttrtcraea	ggaagtgtta	tatgacagra	aaggtaaact	gettgtateg	4020
ccagttgttt	tgttaaacag	aatstaaggt	aaatgacact	ggttaattta	aagtgcatac	4080
atgttgææat	attgcagggc	tatcaagaga	agcacaaaag	gctctaagga	aatatacgtt	4140
gaceggtacc	atttcagtag	gcaagctaac	tgacaattta	acegeteace	gaatgatagg	4200
totottaaac	Attgttaatg	tagatgatgt	ttatgtttca	atctaatagg	gtctacggcg	4260
gtacagctgg	ctccagtacc	acaagtaatg	ggga tacag	atgetge tgg	agttagaast	4320
ataatagaat	ctaacggata	ctttaaaggt	tottorgaat	catggaactg	tgctgttgat	4380
ggtctatatg	aatgtgatgt	tgrecttagt	gctcttggag	agctaattaa	toatetgtet	4440
aggetaaagg	tgtgttggct	tgtttagttt	ttgcttttca	Caaattaagc	aaaggaactt	4500
ttestasett	acagtttcta	totactigca	gctagaagat	gtacttaage	atggggatat	4560
ttttccatac	caagtttaca	ggggttgtct	cagaattgat	ggccagacga	tggtasatct	4620
tgagatattt	Bacaataget	gtgatggtgg	tectteagge	aagcgcatat	ttettttttg	4680
ataacetcaa	ctagagggca	gacatagaag	gaaeeattct	aatacttcgt	acggatetee	4740
agtazgtaat	agcogatttt	tgtttaccta	tgtagggacc	ttgtacaaat	atcttgates	4800
ctgtgttagt	ccaactggta	gegactett	aaggaattgg	atctgocatc	cactcamaga	4860
t gtagaaagc	accaa taaac	ggcttgatgt	agttgaagaa	stcacggcaa	actcagaaag	4920

tatgcaaatc	actggccagt	atctccacaa	acttocagac	ttagaaagac	tàcrcàdeca	4950
catcaagtet	agcgttcgat	catcagecte	tgtgttgcct	getettetgg	ggaaaaaage	5040
gctgaaacaa	cgagtaagta	tcastcacas	gttttctgag	taatgrotto	catgagtagt	5100
acaggactaa	aacattacgg	gtotagotaa	agactgttct	ccttcttttg	caatgtctgg	5160
ttattcatta	catttetett	aacttattgc	attg çaggöt	assgcatttg	ggcaaattgt	5220
gaaagggttc	agzagtggaa	ttgatctgtt	gtrggctcta	cagaaggaat	caaatatgat	5280
gagttegett	tatasactct	gtaaacttee	tatattagta	ggasasagcg	ggctagagtt	5340
artrotttor	caattcgaag	cagccataga	tagegactit	ccaaattatc	aggtgcccat	5400
ctatctttca	tacettacaa	caaaatgtct	gtcactacts	azagezatge	atatggctta	5460
gateteaaet	cacacccccga	ggatcctaaa	gggatttg¢t	tittatteet	aatgtttttg	5520
gatggtttga	tttatttcta	acttgaactt	attaatettg	taccagaacc	aagatgtgac	5580
agatgaaaac	gctgaaactc	tcacaatact	tatcgaactt	cttatogaaa	ga gcaac tea	5640
azggtctgag	gtoattcaca	rcataagctg	cctagatgtc	ctgagatctt	ttgcaatcgc	\$700
agcaagtctc	tctgctggaa	gcatggccag	gcctgttatt	tttcccgaat	cagaagetae	576 0
agatcagaat	cagaaaacaa	aagggccaat	acttaaaatc	caaggactat	ggcatccatt	5820
tgcagttgca	gccgatggtc	aattgeetgt	toogaatgat	atactccttg	gcgaggctag	5880
aagaagtägt	ggcagcattc	atecteggte	attgttactg	асдддассаз	acatāāācāā	5940
assatcaact	cttcttcgtg	caacatgtet	ggccgttatc	tttgcccaag	tttgtatact	6000
cgttagataa	ttactctatt	ctttgcaatc	agttcttcaa	catgaataat	aaattotgtt	6060
trerâterâc	agcttggctg	ctacgtgccg	tgrgagtett	gcgaaatctc	cotogtggat	612(
actatettea	caaggcttgg	cgcatctgat	agaatcatga	caggagagag	taagttttgt	6180
totcaaaata	ccaattcotc	gaactattta	ctcagatttt	gtctgattgg	acaeggtggt	624
tregerttet	tttaggtacc	tttttggtag	aatgcactga	gacagcgtca	gttcttcaga	6300
atgcaactca	ggattcacta	gtaatccttg	acgaactggg	cagaggaact	agtactttcg	6360
arggatacgo	cattgratac	teggtaacet	getettetee	ttcaacttat	acttgttgat	642
CAACRARAEC	atgcaattca	ttttgctgaa	acttattgat	ttatatoagg	tttttcgtca	64B
		~- ~~~ - ~~	655 Been 200	a	ct ct cacca a	6541

a a w c c c a c a	tttcaccesc	grgtcacere	gaaacacatg	gottgogcat	tcaaatcaag	6600
atotgattat	caaccacgtg	grtgtgatca	ag acctagtg	ttettgtadd	gtttaaccga	6560
gggagettgt	ectgagaget	acggacttca	agtggcacte	acggctggaa	taccasacca	6720
agtggttgaa	acagcatcag	gtgctgctca	agccatgaag	agatcaattg	gggaaaactt	6780
caagtcaagt	gagotaagat	ctgagttete	aagtotgcat	gsagactggc	tcaagtcatt	6840
ggtgggta tt	totogagtog	сссасавсаа	tgodcocart	ggcgaagatg	scraegaçaç	6900
titgittigc	ttatggcatg	agatcaeatc	ctcttactgt	gttcccaaat	aaatggctat	696 0
gacataacac	tatctgaagc	togttaagto	ttttgettet	ctgatgttta	ttootottaa	7020
aasatgetta	tatatcaaaa	aattgtttcc	togattataa	caagattata	tatgtatctg	7080
tcggtttagc	tatggtatat	aatatatgta	tgttcatgag	attggtcaag	agaaatactc	7140
acaaacagta	tattäagääg	gaaatatgtt	tatgeattaa	tttaagtttc	aagataaact	7200
gcaaataacc	tcgactadag	ttgcaaagac	CASACSCEES	ttacaaaact	tataagaott	7260
aagttotgaa	ttccctasaa	CCARARRAGA	aaacagaaca	tattttgttg	catctacaaa	7320
caacacaaac	ctacatagtt	tataacttac	tcatcactga	gattaacatc	agaatcattc	7350
tccatttctt	catottoact	ctcatcatca	tcaccaccac	catgatgatt	etectectet	7440
tcacgtaacc	tagcastctc	actitgaget	ctatcaacaa	tetgettett	ctgcaactcc	7500
aaatototet	gaaaatçagc	tctcatcttc	tccaactcct	tcatttgctc	tttcttactc	7560
ttctccatct	tctcataaac	CttCCCAaac	ctitcaacag	aateegeeaa	catcttatac	7620
gaagcagcgt	cattaacett	cttcctctcg	tactozacct	catcatcctc	atcctcctcc	7680
tetteagaat	caccaggact	atccatcatc	toatcaaacc	cattagactt	atctaaataa	7740
accttagtgt	tcatasacac	aaactcacct	gaatcaacac	caczagotaa	acctaaat¢¢	7800
gacttgggcg	aaacacaaag	caacatat c c	aacttattga	aaaacgacca	tttacttgåå	7860
cotaaacctg	atttctcaac	cttaatcttc	tcttttctat	acttcctctt	caagtcatca	7920
atcattctcc	tacattgcgt	ctragatttc	tccatcetta	getecteact	Cactttctca	79 9 0
gdtackteat	tccaatcetc	gttectcaaa	CtCCttctac	ccaattgcaa	aaacctatct	8040
CCCCBAACTT	Caagcaacac	aa				8062

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶: C12N 15/82, 15/29, C07K 14/415, C12N 15/10, 5/04, 5/14, A01H 5/00

(11) International Publication Number:

WO 99/19492

(43) International Publication Date:

22 April 1999 (22.04.99)

(21) International Application Number:

PCT/EP98/06977

(22) International Filing Date:

9 October 1998 (09.10.98)

(30) Priority Data:

PO 9745

10 October 1997 (10.10.97)

ΑU

(71) Applicant (for all designated States except US):
RHONE-POULENC AGRO [FR/FR]; 14/20, rue Pierre
Baizet, F-69009 Lyon (FR).

(72) Inventors; and

- (75) Inventors/Applicants (for US only): DOUTRIAUX, Marie-Pascale [FR/FR]; 64, route de Villebon, F-91160 Saulx les Chartreux (FR). BETZNER, Andreas, Stefan [AU/AU]; 40 Dallachy Place, Page, ACT 2614 (AU). FREYSSINET, Georges [FR/FR]; 21, rue de Nervieux, F-69450 Saint Cyr au Mont d'Or (FR). PEREZ, Pascal [FR/FR]; 17, chemin de la Pradelle, Varennes, F-63450 Chanonat (FR).
- (74) Agent: GENIN, Patrick; Rhône-Poulenc Agro, DPI, 14/20, rue Pierre Baizet, F-69009 Lyon (FR).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
24 June 1999 (24.06.99)

(54) Title: METHODS FOR OBTAINING PLANT VARIETIES

(57) Abstract

An isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide functionally involved in the DNA mismatch repair system of a plant.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES .	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI .	Finland	LT	Lithuania	SK	Słovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
ΑU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	16	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KР	Democratic People's	NZ	New Zealand	2	
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		•
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		
•							

Inte onal Application No PCT/EP 98/06977

a. classification of subject matter IPC 6 C12N15/82 C12N C12N15/29 CO7K14/415 C12N15/10 C12N5/04 C12N5/14 A01H5/00 According to International Patent Classification (IPC) or to both national classification and IPC B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 6 C12N A01H Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practical, search terms used) C. DOCUMENTS CONSIDERED TO BE RELEVANT Category * Citation of document, with indication, where appropriate, of the relevant passages Relevant to claim No. Α WO 95 15381 A (CHAPELLE ALBERT DE ;UNIV 1-5. JOHNS HOPKINS (US)) 8 June 1995 33-35 see the whole document WO 97 37011 A (SETRATECH S A R L ; BORTS 1 - 35RHONA HARRIET (GB); LOUIS EDWARD JOHN (GB) 9 October 1997 see abstract see the whole document Α WO 90 07576 A (SETRATECH) 12 July 1990 1,7-27,31,32 see the whole document and specially page 5, line 16-31, examples 3-4 X Further documents are listed in the continuation of box C. Patent family members are listed in annex. Special categories of cited documents: later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the "A" document defining the general state of the art which is not considered to be of particular relevance invention "E" earlier document but published on or after the international "X" document of particular relevance; the claimed invention filing date cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another document of particular relevance; the claimed invention citation or other special reason (as specified) cannot be considered to involve an inventive step when the document is combined with one or more other such docu-"O" document referring to an oral disclosure, use, exhibition or other means ments, such combination being obvious to a person skilled document published prior to the international filing date but later than the priority date claimed "&" document member of the same patent family Date of the actual completion of the international search Date of mailing of the international search report 16 April 1999 03/05/1999 Name and mailing address of the ISA Authorized officer European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nt.

1

Fax: (+31-70) 340-3016

Mateo Rosell, A.M.

onal Application No. PCT/EP 98/06977 C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT Category Citation of document, with indication, where appropriate, of the relevant passages Relevant to claim No. Α WO 97 01634 A (ANGELETTI P IST RICHERCHE 1.2 BIO ; JIRICNY JOSEF (IT); PALOMBO FABIO () 16 January 1997 see page 1, line 18 - page 2, line 28 see page 58 Α PROLLA T A ET AL: "MLH1, PMS1, AND MSH2 1,2,4,5 INTERACTIONS DURING THE INITATION OF DNA MISMATCH REPAIR IN YEAST" SCIENCE. vol. 265, 19 August 1994, pages 1091-1093, XP000676403 cited in the application Α A. WATANABE ET AL., : "Genomic 1-6 organization and expresssion of the human MSH3 gene" GENOMICS. vol. 31, 1996, pages 311-318, XP002099967 see the whole document and specially Fig.3 ACHARYA S ET AL: "hMSH2 forms specific Α 1-6 mispair-binding complexes with hMSH3 and hMSH6" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA. no. 93, October 1996, pages 13629- 13629, XP002080159 cited in the application sequences from this paper deposited in GenBank under AC= U61981 (hMSH3), and U54777 and U73732-7 (hMSH6). see the whole document K. LIU ET AL., : "characterization of the Α 1-6 mouse Rep-3 gene: sequence similarities to bacterial and yeast mismatch-repair proteins" GENE, vol. 147, 1994, pages 169-177, XP002099968 see the whole document and specially Figure 2. Α I. IACCARINO ET AL., : "MSH6, a 1-6 Saccharomyces cerevisiae protein that binds to mismatches as a heterodimer with MSH2" CURRENT BIOLOGY, vol. 6, no. 4, April 1996, pages 484-486, XP002099969 sequence is deposited at GenBank under AC= Z47746. see the whole document -/--

Inte Jonal Application No PCT/EP 98/06977

CICantinu	Nine) DOGUMENTO CONCERTO TO DE	PCT/EP 9	8/009//
Category '	ation) DOCUMENTS CONSIDERED TO BE RELEVANT Citation of document, with indication, where appropriate, of the relevant passages		Relevant to claim No.
	appropriate, or the relevant passages		Helevant to claim No.
Α	CORRADI A ET AL: "CDNA SEQUENCE, MAP, AND EXPRESSION OF THE MURINE HOMOLOG OF GTBP, ADNA MISMATCH REPAIR GENE" GENOMICS, vol. 36, no. 2, 1 September 1996, pages 288-295, XP000613761 see the whole document and specially Fig.1		1-6
P,X	DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES,13 July 1998, XP002099368 HINXTON, GB AC= AJ007791. Arabidopsis thaliana mRNA for mismatch repair protein (MSH3) see abstract		1-6
P , X	DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES,1 August 1998, XP002099369 HINXTON, GB AC= 065607. Arabidopsis thaliana. Putative mismacht DNA repair protein see abstract		1-5
P,X	DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES,16 October 1997, XP002099371 HINXTON, GB AC= AF009657. Arabidopsis thalian MutS homolog 6-like protein mRNA. see abstract -& K.M. CULLIGAN AND J.B. HAYS: "DNA mismatch repair in plants" PLANT PHYSIOLOGY, vol. 115, 1997, pages 833-839, XP002099372 see the whole document		1-6
	DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES,17 December 1998, XP002099373 HINXTON, GB AC= AJ131669. Triticum aestivum mRNA for MSH3 protein, MSH3 gene. see abstract		1-6
	DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES,12 October 1998, XP002099370 HINXTON, GB AC= AJ007792. Arabidopsis thaliana DNA mismatch repair protein, MSH6 gene. see abstract		1-6

1 .

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

	Info	Information on patent family members				P 98/06977
Patent document cited in search report		Publication date		Patent family member(s)		Publication date
WO 9515381	A	08-06-1995	EP JP US US	073064 950650 569347 583744	9 T 0 A	11-09-1996 30-06-1997 02-12-1997 17-11-1998
WO 9737011	Α	09-10-1997	NONE	-		
WO 9007576	A	12-07-1990	FR AT AU CA DE DE EP ES IE JP	264179 12751 483439 200654 6892417 6892417 044992 207705 7246 450360	9 T 00 A 9 A 24 D 24 T 23 A 8 T 9 B	20-07-1990 15-09-1995 01-08-1990 26-06-1990 12-10-1995 14-03-1996 09-10-1991 16-11-1995 09-04-1997 02-07-1992
WO 9701634	Α	16-01-1997	IT AU	RM95043 624129		27-12-1996 30-01-1997